

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:53:11 ; Search time 28.875 Seconds  
(without alignments)  
49.473 Million cell updates/sec

Title: US-09-228-866-1  
Perfect score: 9  
Sequence: 1 CNSRLHLRC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 358712

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID		Description
1	9	100.0	9	18	AAW13410		Brain homing pepti
2	9	100.0	9	21	AAB07387		Brain homing pepti
3	9	100.0	9	22	AAE11793		Phage peptide #1 t
4	9	100.0	9	23	AAU10704		Brain homing pepti
5	9	100.0	9	24	ABU59529		Brain receptor tar
6	5	55.6	9	18	AAW13411		Brain homing pepti
7	5	55.6	9	21	AAB07391		Brain homing pepti
8	5	55.6	9	22	AAE11797		Phage peptide #5 t
9	5	55.6	9	23	AAU10708		Brain homing pepti
10	5	55.6	11	24	ABJ36846		G protein coupled
11	5	55.6	21	23	AAU89223		Insulin/insulin-li
12	4	44.4	7	20	AAU16941		Heat shock protein
13	4	44.4	7	21	AAU94221		Murine 16E10 light
14	4	44.4	7	22	AAU72083		Melanoma antigen,
15	4	44.4	7	22	ABB55960		Vascular dementia-
16	4	44.4	7	22	AAM46925		H11 binding site c
17	4	44.4	7	22	AAM46930		H11 binding site c
18	4	44.4	7	22	AAM46965		H11 binding site c
19	4	44.4	7	23	AAU80633		Javelin peptide #6
20	4	44.4	7	23	AAU70359		Mouse Kappa V ligh
21	4	44.4	8	20	AAU48874		Membrane dipeptida
22	4	44.4	8	24	ABJ18554		Ganglioside-associ
23	4	44.4	9	20	AAU48679		Membrane dipeptida
24	4	44.4	9	22	AAU06377		Human Leukocyte An
25	4	44.4	9	23	ABJ04445		Stem cell (mesench
26	4	44.4	9	23	ABJ04447		Stem cell (mesench
27	4	44.4	9	23	AAE22284		Murine normal SPAS
28	4	44.4	9	23	AAU92289		PHOR1-F5D6 peptide
29	4	44.4	9	23	AAU92307		PHOR1-F5D6 peptide
30	4	44.4	9	23	AAU92870		PHOR1-F5D6 peptide
31	4	44.4	9	23	ABB05278		Vascular endotheli
32	4	44.4	9	24	ABR07510		Human cancer-relat
33	4	44.4	9	24	ABR07546		Human cancer-relat
34	4	44.4	9	24	ABR07706		Human cancer-relat
35	4	44.4	9	24	ABR07714		Human cancer-relat
36	4	44.4	9	24	ABR07923		Human cancer-relat
37	4	44.4	9	24	ABR08123		Human cancer-relat
38	4	44.4	9	24	ABR08333		Human cancer-relat
39	4	44.4	9	24	ABR08364		Human cancer-relat
40	4	44.4	9	24	ABR08499		Human cancer-relat
41	4	44.4	9	24	ABR08527		Human cancer-relat
42	4	44.4	9	24	ABR08540		Human cancer-relat
43	4	44.4	9	24	ABR08564		Human cancer-relat
44	4	44.4	9	24	ABR08708		Human cancer-relat
45	4	44.4	9	24	ABR08744		Human cancer-relat

# ALIGNMENTS

RESULT 1

AAW13410

ID AAW13410 standard; Peptide; 9 AA.

XX

AC AAW13410;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 1997-202359/18.

XX

PT Obtaining compound that homes to selected organ or tissue - by in  
PT vivo panning method, specifically to identify brain, kidney,  
PT angiogenic vasculature or tumour tissue homing peptide(s)

XX

PS Claim 11; Page 67; 75pp; English.

XX

CC This synthetic peptide is a claimed example of a brain-homing  
CC peptide that was identified using a novel method for obtaining  
CC molecules that home to a selected organ or tissue. This in vivo  
CC panning method typically involves administering a phage display  
CC library to a subject, and identifying expressed peptides which  
CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
CC vascular tissue or tumour tissue. The isolated peptides (see  
CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
CC labels to the selected organ/tissue (claimed) or to identify and/or  
CC isolate target molecules (claimed). The peptides can be directly  
CC identified in vivo, as compared to prior art in vitro screening  
CC methods, which require further examination to see if they maintain  
CC specificity in vivo.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 18; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRLHLRC 9

|||||||

Db 1 CNSRLHLRC 9

RESULT 2

AAB07387

ID AAB07387 standard; peptide; 9 AA.

XX

AC AAB07387;

XX

DT 17-OCT-2000 (first entry)

XX

DE Brain homing peptide # 1.

XX

KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS Mus sp.

XX

FH Key Location/Qualifiers

FT Disulfide-bond 1..9

FT /note= "Can optionally form a cyclic peptide"

XX

PN US6068829-A.

XX

PD 30-MAY-2000.

XX

PF 23-JUN-1997; 97US-0862855.

XX

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 2000-410850/35.

XX

PT Identifying and recovering organ homing molecules or peptides by in

PT vivo panning comprises administering a library of diverse peptides

PT linked to a tag which facilitates recovery of these peptides -

XX

PS Example 2; Column 17; 20pp; English.

XX

CC The present sequence is a mouse brain homing peptide. This sequence was

CC identified by using in vivo panning to screen a library of potential

CC organ homing molecules. The present sequence can be used to direct a

CC moiety to a the brain tissue, by linking the moiety to the present

CC sequence. Examples of potential moieties are drugs, toxins or a

CC detectable label. The present sequence contains a SRL amino acid motif.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 21; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRLHLRC 9

|||||||



Db

1 CNSRLHLRC 9

RESULT 3

AAE11793

ID AAE11793 standard; peptide; 9 AA.

XX

AC AAE11793;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #1 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;  
KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

FH Key Location/Qualifiers

FT Domain 3..5

FT /label= SRL\_motif

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo  
PT panning that selectively home to a selected organ or tissue useful for  
PT treating disease or in diagnostic methods -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 22; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRLHLRC 9  
|||  
Db 1 CNSRLHLRC 9

RESULT 4

AAU10704

ID AAU10704 standard; peptide; 9 AA.

XX

AC AAU10704;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #1 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;

KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
PT identifying molecules that home to a specific organ or tissue, e.g.  
PT identifying a tumour homing molecule to identify the presence of cancer,  
PT by in vivo panning of a library -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
CC home to a selected organ or tissue. The method comprises administering  
CC to the subject the library of diverse molecules, collecting a sample of  
CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
CC the sample several molecules that home to the selected organ or tissue.  
CC The method is useful for identifying molecules, particularly useful for  
CC screening large number of molecules (e.g. peptides), that home to a  
CC specific organ. The identified molecule is useful for e.g. raising an  
CC antibody specific for a target molecule, targeting a desired moiety  
CC (e.g. drug, toxin or detectable label) to the selected organ.

CC Specifically, the method is useful for identifying the presence of cancer  
CC in a subject by linking an appropriate moiety to a tumour homing  
CC molecule. The present method provides a direct means for identifying  
CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 23; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRLHLRC 9

|||||||

Db 1 CNSRLHLRC 9

RESULT 5

ABU59529

ID ABU59529 standard; Peptide; 9 AA.

XX

AC ABU59529;

XX

DT 22-APR-2003 (first entry)

XX

DE Brain receptor targeting peptide #1.

XX

KW Targeting ligand; bioactive agent; polymer matrix; cancer; cytostatic;

KW cathepsin-D substrate; peptides; neuroreceptor; adrenal receptor;

KW fibronectin; vitronectin; integrin; RGD motif; angiogenic endothelium;

KW tumour; cationic cancer-targeting peptide.

XX

OS Synthetic.

XX

PN ~~US2002011898-A1~~

US 20020041898

XX

PD 11-APR-2002.

XX

PF 25-JUL-2001; 2001US-0912609.

XX

PR 05-JAN-2000; 2000US-0478124.

PR 31-OCT-2000; 2000US-0703474.

XX

PA (UNGE/) UNGER E C.

PA (MATS/) MATSUNAGA T O.

PA (RAMA/) RAMASWAMI V.

PA (ROMA/) ROMANOWSKI M J.

XX

PI Unger EC, Matsunaga TO, Ramaswami V, Romanowski MJ;

XX

DR WPI; 2003-208921/20.

XX

PT Targeted delivery system comprising a bioactive agent homogeneously

PT dispersed in a targeted matrix is especially useful in cancer therapy  
PT -

XX

PS Claim 23; Page 37; 46pp; English.

XX

CC The invention relates to a composition comprising a bioactive agent  
CC homogeneously dispersed in a targeted matrix (polymer and targeting  
CC ligand). Also included are a targeted matrix for use as a delivery  
CC vehicle comprising a polymer associated with a targeting ligand,  
CC enhancing the bioavailability of an agent comprising administration  
CC of the composition and treating cancer comprising administration of the  
CC novel composition. The method is useful for targeted delivery of a drug,  
CC especially in cancer therapy. The targeting ligand may be a peptide.  
CC Examples of targeting peptides are disclosed including cathepsin-D  
CC substrate peptides, peptides targeting receptors in the brain and  
CC kidney, peptides recognising fibronectin- and vitronectin-binding  
CC integrins, peptides targeting the RGD (Arg-Gly-Asp)-motif in, e.g.,  
CC antibodies, peptides targeting the angiogenic endothelium of solid  
CC tumours, tissue specific peptides (e.g. of lung, skin, pancreas,  
CC intestine, uterus, adrenal gland and retina), and cationic cancer-  
CC targeting peptides. The present sequence is a peptide targeting  
CC ligand disclosed in the invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 24; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRLHLRC 9  
|||  
Db 1 CNSRLHLRC 9

#### RESULT 6

AAW13411

ID AAW13411 standard; Peptide; 9 AA.

XX

AC AAW13411;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;  
KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX  
PA (IJOL-) LA JOLLA CANCER RES FOUND.  
XX  
PI Pasqualini R, Ruoslahti E;  
XX  
DR WPI; 1997-202359/18.  
XX  
PT Obtaining compound that homes to selected organ or tissue - by in  
PT vivo panning method, specifically to identify brain, kidney,  
PT angiogenic vasculature or tumour tissue homing peptide(s)  
XX  
PS Claim 11; Page 67; 75pp; English.  
XX  
CC This synthetic peptide is a claimed example of a brain-homing  
CC peptide that was identified using a novel method for obtaining  
CC molecules that home to a selected organ or tissue. This in vivo  
CC panning method typically involves administering a phage display  
CC library to a subject, and identifying expressed peptides which  
CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
CC vascular tissue or tumour tissue. The isolated peptides (see  
CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
CC labels to the selected organ/tissue (claimed) or to identify and/or  
CC isolate target molecules (claimed). The peptides can be directly  
CC identified in vivo, as compared to prior art in vitro screening  
CC methods, which require further examination to see if they maintain  
CC specificity in vivo.  
XX  
SQ Sequence 9 AA;

Query Match 55.6%; Score 5; DB 18; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRL 5  
|||||  
Db 1 CNSRL 5

RESULT 7  
AAB07391  
ID AAB07391 standard; peptide; 9 AA.  
XX  
AC AAB07391;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Brain homing peptide # 5.  
XX  
KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..9  
FT /note= "Can optionally form a cyclic peptide"  
XX  
PN US6068829-A.

XX  
PD 30-MAY-2000.  
XX  
PF 23-JUN-1997; 97US-0862855.  
XX  
PR 11-SEP-1995; 95US-0526710.  
PR 10-MAR-1997; 97US-0813273.  
XX  
PA (BURN-) BURNHAM INST.  
XX  
PI Pasqualini R, Ruoslahti E;  
XX  
DR WPI; 2000-410850/35.  
XX  
PT Identifying and recovering organ homing molecules or peptides by in  
PT vivo panning comprises administering a library of diverse peptides  
PT linked to a tag which facilitates recovery of these peptides -  
XX  
PS Example 2; Column 17; 20pp; English.  
XX  
CC The present sequence is a mouse brain homing peptide. This sequence was  
CC identified by using in vivo panning to screen a library of potential  
CC organ homing molecules. The present sequence can be used to direct a  
CC moiety to a the brain tissue, by linking the moiety to the present  
CC sequence. Examples of potential moieties are drugs, toxins or a  
CC detectable label. The present sequence contains a SRL amino acid motif.  
XX  
SQ Sequence 9 AA;

Query Match 55.6%; Score 5; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRL 5  
|||||  
Db 1 CNSRL 5

RESULT 8  
AAE11797  
ID AAE11797 standard; peptide; 9 AA.  
XX  
AC AAE11797;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Phage peptide #5 targetted to brain.  
XX  
KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;  
KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.  
XX  
OS Bacteriophage.  
XX  
FH Key Location/Qualifiers  
FT Domain 3..5  
FT /label= SRL\_motif  
XX

PN US6296832-B1.  
 XX  
 PD 02-OCT-2001.  
 XX  
 PF 08-JAN-1999; 99US-0226985.  
 XX  
 PR 23-JUN-1997; 97US-0862855.  
 PR 11-SEP-1995; 95US-0526710.  
 PR 10-MAR-1997; 97US-0813273.  
 XX  
 PA (BURN-) BURNHAM INST.  
 XX  
 PI Ruoslahti E, Pasqualini R;  
 XX  
 DR WPI; 2001-610691/70.  
 XX  
 PT Enriched library fraction comprising molecules recovered by in vivo  
 PT panning that selectively home to a selected organ or tissue useful for  
 PT treating disease or in diagnostic methods -  
 XX  
 PS Example 2; Column 17; 21pp; English.  
 XX  
 CC The invention relates to an enriched library fraction containing  
 CC molecules that selectively home to a selected organ or tissue such as  
 CC brain, kidney or tumour recovered by in vivo panning. The invention  
 CC generally relates to the field of molecular medicine, drug delivery and  
 CC to a method of invivo panning for identifying a molecule that homes to a  
 CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
 CC and fragments of proteins contained in an enriched library fraction may  
 CC be administered to a subject as part of a pharmaceutical composition to  
 CC treat disease or in diagnostic methods. The present sequence is a  
 CC peptide from bacteriophage targetted to brain.  
 XX  
 SQ Sequence 9 AA;  
  
 Query Match 55.6%; Score 5; DB 22; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSRL 5  
 |||||  
 Db 1 CNSRL 5

RESULT 9  
 AAU10708  
 ID AAU10708 standard; peptide; 9 AA.  
 XX  
 AC AAU10708;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Brain homing peptide #5 useful for delivery of target molecules.  
 XX  
 KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
 KW delivery of target molecule; brain homing peptide.  
 XX

OS Synthetic.  
 XX  
 PN US6306365-B1.  
 XX  
 PD 23-OCT-2001.  
 XX  
 PF 08-JAN-1999; 99US-0227906.  
 XX  
 PR 23-JUN-1997; 97US-0862855.  
 PR 11-SEP-1995; 95US-0526710.  
 PR 10-MAR-1997; 97US-0813273.  
 XX  
 PA (BURN-) BURNHAM INST.  
 XX  
 PI Ruoslahti E, Pasqualini R;  
 XX  
 DR WPI; 2002-040196/05.  
 XX  
 PT Recovering molecules that home to an organ or tissue, useful for  
 PT identifying molecules that home to a specific organ or tissue, e.g.  
 PT identifying a tumour homing molecule to identify the presence of cancer,  
 PT by in vivo panning of a library -  
 XX  
 PS Example 2; Column 17; 21pp; English.  
 XX  
 CC The present invention relates to a method of recovering molecules that  
 CC home to a selected organ or tissue. The method comprises administering  
 CC to the subject the library of diverse molecules, collecting a sample of  
 CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
 CC the sample several molecules that home to the selected organ or tissue.  
 CC The method is useful for identifying molecules, particularly useful for  
 CC screening large number of molecules (e.g. peptides), that home to a  
 CC specific organ. The identified molecule is useful for e.g. raising an  
 CC antibody specific for a target molecule, targeting a desired moiety  
 CC (e.g. drug, toxin or detectable label) to the selected organ.  
 CC Specifically, the method is useful for identifying the presence of cancer  
 CC in a subject by linking an appropriate moiety to a tumour homing  
 CC molecule. The present method provides a direct means for identifying  
 CC molecules that specifically home to a selected organ and, therefore  
 CC provides a significant advantage over previous methods, which require  
 CC that a molecule identified using an in vitro screening method  
 CC subsequently be examined to determine if it maintains its specificity in  
 CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
 CC the present invention.  
 XX  
 SQ Sequence 9 AA;  
  
 Query Match 55.6%; Score 5; DB 23; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 1 CNSRL 5  
 |||||  
 Db 1 CNSRL 5

RESULT 10



ABJ36846

ID ABJ36846 standard; Peptide; 11 AA.

XX

AC ABJ36846;

XX

DT 01-MAY-2003 (first entry)

XX

DE G protein coupled receptor related peptide SEQ ID No 200.

XX

KW Nootropic; cardiant; antiarteriosclerotic; hypotensive; cytostatic;  
KW antibacterial; analgesic; antiallergic; antiasthmatic; antiinflammatory;  
KW osteopathic; neuroprotective; anxiolytic; anorectic; lead compound;  
KW G protein coupled receptor signaling inhibitor; GPCR; library;  
KW high throughput screening assay; stroke; myocardial infarction;  
KW restenosis; atherosclerosis; hypotension; cancer; infection; asthma;  
KW septic shock; pain; allergic disorder; inflammatory bowel disease;  
KW osteoporosis; obesity; psychotic; neurological disorder; anxiety;  
KW schizophrenia; Alzheimer's disease.

XX

OS Unidentified.

XX

PN WO200272778-A2.

XX

PD 19-SEP-2002.

XX

PF 14-MAR-2002; 2002WO-US07561.

XX

PR 14-MAR-2001; 2001US-275472P.

PR 11-MAY-2001; 2001US-0852910.

XX

PA (CUEB-) CUE BIOTECH.

XX

PI Gilchrist A, Hamm HE;

XX

DR WPI; 2003-247841/24.

XX

PT Identifying G protein coupled receptor (GPCR) signaling inhibitors,  
PT useful in screening drugs for treating stroke, cancers or pain, by  
PT identifying compounds that block GPCR mediated signaling with high  
PT affinity and specificity -

XX

PS Claim 94; Page 66; 94pp; English.

XX

CC The invention relates to a novel method for identifying a G protein  
CC coupled receptor (GPCR) signaling inhibitor. The novel method comprises  
CC selecting or identifying a member of a library of peptides and/or  
CC candidate compounds, having binding to a GPCR of higher affinity than  
CC that of the native peptide. The peptide library is based on a native GPCR  
CC binding peptide. The method is useful for identifying inhibitors of a G  
CC protein coupled receptor (GPCR) signaling. The method is particularly  
CC useful for identifying drugs that antagonise the binding between a GPCR  
CC and its extracellular ligand(s). The method is especially useful in  
CC modern high throughput screening assays for identifying potent lead  
CC compounds. The compounds, peptides or inhibitors identified by the method  
CC are useful for preventing, ameliorating or treating diseases in which  
CC GPCR signaling is a causative factor or in which a specific class of G  
CC protein is relevant, e.g. stroke, myocardial infarction, restenosis,

CC atherosclerosis, hypotension, cancers, infections, septic shock, pain,  
CC allergic disorders, asthma, inflammatory bowel disease, osteoporosis,  
CC obesity, or psychotic and neurological disorders (e.g. anxiety,  
CC schizophrenia or Alzheimer's disease). This sequence represents a peptide  
CC relating to the G protein coupled receptors of the invention.

XX

SQ Sequence 11 AA;

Query Match 55.6%; Score 5; DB 24; Length 11;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLHLR 8

|||||

Db 2 RLHLR 6

#### RESULT 11

AAU89223

ID AAU89223 standard; Peptide; 21 AA.

XX

AC AAU89223;

XX

DT 18-JUN-2002 (first entry)

XX

DE Insulin/insulin-like growth factor receptor-binding peptide #1179.

XX

KW Cytostatic; antidiabetic; neuroprotective; cerebroprotective;

KW ophthalmological; insulin; receptor; gene therapy; diabetes;

KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;

KW diabetic retinopathy; neurological diseases; stroke;

KW diabetic neuropathy.

XX

OS Synthetic.

XX

PN WO200172771-A2.

XX

PD 04-OCT-2001.

XX

PF 29-MAR-2000; 2000WO-US08528.

XX

PR 29-MAR-2000; 2000WO-US08528.

XX

PA (DGIB-) DGI BIOTECHNOLOGIES LLC.

PA (NOVO ) NOVO NORDISK AS.

XX

PI Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;

PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandecki WS;

PI Hansen PH, Ravera M, Hsiao K;

XX

DR WPI; 2002-025774/03.

XX

PT Modulating insulin activity in mammalian cells, for treating e.g.

PT diabetes and tumours, comprises using peptides that bind to insulin or

PT insulin-like growth factor receptors -

XX

PS Disclosure; Figure 10-1; 390pp; English.

XX  
 CC The invention relates to a method of modulating insulin activity in  
 CC mammalian cells by administering a peptide that binds the insulin  
 CC receptor (IR). A composition containing a peptide, optionally expressed  
 CC from gene therapy vectors, that binds to Site 1 of IR and an insulin  
 CC agonist are useful for treating diabetes. Also, peptides that are  
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases,  
 CC including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AAU88034-AAU90957 represent IR  
 CC and/or IGF-1 receptor-binding peptides and related amino acid sequences  
 CC of the invention.  
 XX  
 SQ Sequence 21 AA;

Query Match 55.6%; Score 5; DB 23; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLHLR 8  
 |||||  
 Db 2 RLHLR 6

RESULT 12  
 AAY16941  
 ID AAY16941 standard; peptide; 7 AA.  
 XX  
 AC AAY16941;  
 XX  
 DT 20-JUL-1999 (first entry)  
 XX  
 DE Heat shock protein (hsp) binding peptide.  
 XX  
 KW Conjugate peptide; heat shock protein; hsp; phage display library; virus;  
 KW surface protein; tethering peptide; chaperone process; cytokine; cancer;  
 KW neoplastic disease; infectious disease; bacterium; immune system; fungus;  
 KW acquired immune deficiency; autoimmune disease.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9922761-A1.  
 XX  
 PD 14-MAY-1999.  
 XX  
 PF 22-OCT-1998; 98WO-US22335.  
 XX  
 PR 31-OCT-1997; 97US-0961707.  
 XX  
 PA (SLOK ) SLOAN KETTERING INST CANCER RES.  
 XX  
 PI Hartl U, Hoe MH, Houghton A, Mayhew M, Moroi Y;  
 PI Ouerfelli O, Rothman JE;  
 XX

DR WPI; 1999-313177/26.  
 XX  
 PT Identifying peptides which bind heat shock proteins  
 XX  
 PS Examples; Page 22; 155pp; English.  
 XX  
 CC The invention relates to conjugate peptides engineered to noncovalently  
 CC bind to heat shock proteins (hsp). A method of identifying a hsp binding  
 CC peptide comprises (a) contacting a phage display library having  
 CC bacteriophage expressing, in a surface protein, inserted peptides with a  
 CC hsp target, and bound to a benzoquinone ansamycin antibiotic (BAA), in a  
 CC physiologic binding buffer; (b) isolating a phage binding to the hsp  
 CC target; and (c) identifying the inserted peptide expressed. The peptides  
 CC which bind to a hsp can be used as tethering peptides for a hsp which may  
 CC serve as an accessory in a chaperone process and/or may comprise a  
 CC cytokine. They can also be coupled to antigens to induce an immune  
 CC response. Such compositions can be used for treating neoplastic disease,  
 CC e.g. cancers, infectious diseases, e.g. diseases caused by a bacterium,  
 CC virus, protozoan, mycoplasma, fungus, yeast, parasite or prion, or a  
 CC disease of the immune system, e.g. acquired immune deficiencies or  
 CC autoimmune diseases.  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 20; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSRL 5  
 ||||  
 Db 1 NSRL 4

# RESULT 13

AAAY94221

ID AAY94221 standard; Peptide; 7 AA.

XX

AC AAY94221;

XX

DT 08-AUG-2000 (first entry)

XX

DE Murine 16E10 light chain complementarily determining region 2.

XX

KW Antibody; RHAMM; receptor for hyaluronic acid mediated motility;  
 KW ras-dependent proliferation; leukaemia; cancer; lymphoma;  
 KW inflammatory disease; proliferative disease; psoriasis;  
 KW inflammatory bowel disease; rheumatoid arthritis;  
 KW proliferative cardiovascular disease; restenosis;  
 KW proliferative ocular disorder; diabetic retinopathy; haemangioma;  
 KW benign hyperproliferative disease; tumour formation; light chain;  
 KW variable region; 16E10; CDR2; complementarily determining region 2.

XX

OS Mus musculus.

XX

PN WO200029447-A1.

XX

PD 25-MAY-2000.

XX  
 PF 19-NOV-1999; 99WO-US27565.  
 XX  
 PR 19-NOV-1998; 98US-0109041.  
 PR 14-JUL-1999; 99US-0143692.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Abrahamson JA, Holmes SD, Jackson JR;  
 XX  
 DR WPI; 2000-387752/33.  
 XX  
 PT Antibodies against receptor for hyaluronic acid mediated motility,  
 PT useful for treating or preventing proliferative diseases, e.g. cancer  
 PT or cardiovascular disease -  
 XX  
 PS Claim 23; Page 35; 39pp; English.  
 XX  
 CC The present sequence is the amino acid sequence for the murine 16E10  
 CC light chain complementarily determining region 2. This forms part of the  
 CC monoclonal antibody 16E10, which can be used against the receptor for  
 CC hyaluronic acid mediated motility (RHAMM). RHAMM is required for  
 CC ras-transformation of cells, which leads to tumour formation, and so the  
 CC antibody can be used to treat proliferative disorders such as leukaemias,  
 CC solid tumour cancers and metastases including lymphomas, soft tissue,  
 CC brain, oesophageal, stomach, pancreatic, liver, lung, bladder, bone,  
 CC prostate, ovarian, cervical, uterine, skin, breast, testicular, kidney,  
 CC head and neck and colon cancers, chronic inflammatory diseases such as  
 CC psoriasis, inflammatory bowel disease and rheumatoid arthritis,  
 CC proliferative cardiovascular diseases such as restenosis, proliferative  
 CC ocular disorders such as diabetic retinopathy and benign  
 CC hyperproliferative diseases such as haemangiomas.  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 21; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRLH 6  
 ||||  
 Db 3 SRLH 6

#### RESULT 14

AAU72083

ID AAU72083 standard; Peptide; 7 AA.

XX

AC AAU72083;

XX

DT 26-FEB-2002 (first entry)

XX

DE Melanoma antigen, javelin peptide #69.

XX

KW Melanoma antigen; MART-1; MAGE-1; gp100; cytostatic; immune response;

KW immunotherapeutic; heat shock protein; tyrosinase; BAGE; NYEs01; GM2;

KW tyrosinase related protein 1; tyrosinase related protein 2; vaccine;  
 KW javelin molecule; melanoma antigen recognised by T cells-1; human.  
 XX  
 OS Bacteriophage M13.  
 XX  
 PN WO200178655-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 17-APR-2001; 2001WO-US12449.  
 XX  
 PR 17-APR-2000; 2000US-197462P.  
 XX  
 PA (HOUG/) HOUGHTON A.  
 PA (LIVI/) LIVINGSTON P.  
 PA (ALAW/) AL-AWQATI Q.  
 PA (MAYH/) MAYHEW M.  
 PA (HOEM/) HOE M.  
 XX  
 PI Houghton A, Livingston P, Al-awqati Q, Mayhew M, Hoe M;  
 XX  
 DR WPI; 2001-663092/76.  
 XX  
 PT Anti cancer vaccine for the treatment of melanoma comprises a heat  
 PT shock protein and a melanoma antigen i.e. tyrosinase -  
 XX  
 PS Disclosure; Page 17; 150pp; English.  
 XX  
 CC The invention relates to a method of induction of an immune response,  
 CC comprising administration of an immunotherapeutic composition, comprising  
 CC a heat shock protein, and a melanoma antigen, where the melanoma  
 CC antigen is selected from tyrosinase, tyrosinase related protein 1,  
 CC tyrosinase related protein 2, gp 100, MAGE antigens, BAGE antigens,  
 CC NYEs01, MART antigens, GM2, antigenic portions and combinations of these.  
 CC The melanoma antigen is covalently bound to a javelin molecule, where the  
 CC melanoma antigen bound to the javelin molecule is non-covalently bound to  
 CC the heat shock protein. The composition is useful for inducing an immune  
 CC response for the treatment of melanoma. AAU71980-AAU72481 represent  
 CC melanoma antigen peptides of the invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSRL 5  
 ||||  
 Db 1 NSRL 4

RESULT 15  
 ABB55960  
 ID ABB55960 standard; Peptide; 7 AA.  
 XX  
 AC ABB55960;  
 XX

DT 15-FEB-2002 (first entry)  
 XX  
 DE Vascular dementia-associated protein isoform (VPI) 160.  
 XX  
 KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;  
 KW diagnosis; prognosis; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200169261-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 14-MAR-2001; 2001WO-GB01106.  
 XX  
 PR 15-MAR-2000; 2000GB-0006285.  
 PR 24-NOV-2000; 2000GB-0028734.  
 PR 28-NOV-2000; 2000US-0724391.  
 XX  
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX  
 PI Herath HMAc, Parekh RB, Rohlf C;  
 XX  
 DR WPI; 2001-557937/62.  
 XX  
 PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for  
 PT determining stage of VD and monitoring the effect of VD therapy,  
 PT comprises analysing body fluid by 2-dimensional electrophoresis for  
 PT features correlated with VD -  
 XX  
 PS Claim 6; Page 33; 151pp; English.  
 XX  
 CC The invention relates to screening, diagnosis or prognosis of Vascular  
 CC Dementia (VD) in a subject comprising analysing body fluid from the  
 CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of  
 CC features containing at least one chosen feature whose relative abundance  
 CC correlates with the presence, absence, stage or severity of VD or  
 CC predicts the onset or course of VD, especially detecting in a sample of  
 CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated  
 CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the  
 CC specification. Detecting VD-associated features and VPI is useful for the  
 CC screening, diagnosis or prognosis of VD, for determining the stage or  
 CC severity of VD, for identifying a subject at risk of VD or for  
 CC monitoring the effect of therapy administered to a subject having VD.  
 CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are  
 CC useful for the treatment of VD and for gene therapy.  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LHLR 8  
 ||||  
 Db 4 LHLR 7

Search completed: November 13, 2003, 10:32:51  
Job time : 29.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:38:27 ; Search time 17.625 Seconds  
(without alignments)  
93.222 Million cell updates/sec

Title: US-09-228-866-1  
Perfect score: 9  
Sequence: 1 CNSRLHLRC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 124183

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES



Result No.	Score	% Query Match	Length	DB	ID	Description
1	9	100.0	9	12	US-10-306-878-11	Sequence 11, Appl
2	5	55.6	9	15	US-10-158-596A-107	Sequence 107, App
3	5	55.6	9	15	US-10-254-446A-234	Sequence 234, App
4	5	55.6	11	11	US-09-852-910-200	Sequence 200, App
5	5	55.6	18	12	US-10-029-386-27913	Sequence 27913, A
6	5	55.6	21	12	US-09-962-756-677	Sequence 677, App
7	4	44.4	7	11	US-09-563-222-39	Sequence 39, Appl
8	4	44.4	7	12	US-10-052-578-211	Sequence 211, App
9	4	44.4	7	12	US-10-053-520-211	Sequence 211, App
10	4	44.4	7	12	US-10-053-498B-211	Sequence 211, App
11	4	44.4	7	14	US-10-140-555-9	Sequence 9, Appli
12	4	44.4	7	15	US-10-139-496-11	Sequence 11, Appl
13	4	44.4	9	9	US-09-832-723-110	Sequence 110, App
14	4	44.4	9	10	US-09-952-432A-23	Sequence 23, Appl
15	4	44.4	9	12	US-10-303-331-110	Sequence 110, App
16	4	44.4	10	10	US-09-952-432A-31	Sequence 31, Appl
17	4	44.4	11	11	US-09-852-910-69	Sequence 69, Appl
18	4	44.4	11	12	US-10-137-867-428	Sequence 428, App
19	4	44.4	12	11	US-09-932-613-20	Sequence 20, Appl
20	4	44.4	12	12	US-10-158-825-100	Sequence 100, App
21	4	44.4	12	12	US-10-286-457-355	Sequence 355, App
22	4	44.4	12	12	US-09-932-322-20	Sequence 20, Appl
23	4	44.4	12	15	US-10-158-847-100	Sequence 100, App
24	4	44.4	12	15	US-10-254-446A-213	Sequence 213, App
25	4	44.4	13	10	US-09-851-138-156	Sequence 156, App
26	4	44.4	13	12	US-10-268-332-47	Sequence 47, Appl
27	4	44.4	13	12	US-09-964-821B-47	Sequence 47, Appl
28	4	44.4	13	15	US-10-097-065-454	Sequence 454, App
29	4	44.4	17	9	US-09-864-675-17	Sequence 17, Appl
30	4	44.4	17	12	US-10-299-003-14	Sequence 14, Appl
31	4	44.4	17	12	US-09-962-756-1642	Sequence 1642, Ap
32	4	44.4	18	9	US-09-836-861-9	Sequence 9, Appli
33	4	44.4	18	15	US-10-084-813-132	Sequence 132, App
34	4	44.4	18	15	US-10-084-813-133	Sequence 133, App
35	4	44.4	18	15	US-10-084-813-134	Sequence 134, App
36	4	44.4	18	15	US-10-084-813-135	Sequence 135, App
37	4	44.4	18	15	US-10-225-567A-865	Sequence 865, App
38	4	44.4	19	9	US-09-853-080-27	Sequence 27, Appl
39	4	44.4	19	12	US-09-962-756-1213	Sequence 1213, Ap
40	4	44.4	20	12	US-10-340-288-5	Sequence 5, Appli
41	4	44.4	20	12	US-10-280-066-153	Sequence 153, App
42	4	44.4	20	15	US-10-225-567A-1789	Sequence 1789, Ap
43	4	44.4	20	15	US-10-225-567A-1799	Sequence 1799, Ap
44	4	44.4	21	9	US-09-864-761-44104	Sequence 44104, A
45	4	44.4	21	10	US-09-842-256-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-10-306-878-11  
; Sequence 11, Application US/10306878

```
; Publication No. US20030175819A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Guo, Bin
; TITLE OF INVENTION: Methods for Identifying Modulators of
; TITLE OF INVENTION: Apoptosis
; FILE REFERENCE: P-LJ 5535
; CURRENT APPLICATION NUMBER: US/10/306,878
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: US 60/334,149
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-10-306-878-11
```

```
Query Match          100.0%; Score 9; DB 12; Length 9;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      '9; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

```
Qy      1 CNSRLHLRC 9
        |||||
Db      1 CNSRLHLRC 9
```

## RESULT 2

```
US-10-158-596A-107
; Sequence 107, Application US/10158596A
; Publication No. US20030068900A1
; GENERAL INFORMATION:
; APPLICANT: Belcher, Angela
; APPLICANT: Flynn, Christine
; TITLE OF INVENTION: BIOLOGICAL CONTROL OF NANOPARTICLE NUCLEATION, SHAPE AND
CRYSTAL PHASE
; FILE REFERENCE: 119927-1052
; CURRENT APPLICATION NUMBER: US/10/158,596A
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/296,013
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 107
; LENGTH: 9
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: peptide retrieved from phage biopanning
US-10-158-596A-107
```

```
Query Match          55.6%; Score 5; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      5; Conservative    0; Mismatches    0; Indels      0; Gaps      0;
```

Qy 3 SRLHL 7  
 |||||  
Db 4 SRLHL 8

RESULT 3

US-10-254-446A-234

; Sequence 234, Application US/10254446A  
; Publication No. US20030113714A1  
; GENERAL INFORMATION:  
; APPLICANT: Belcher, Angela M  
; APPLICANT: Smalley, Richard E.  
; APPLICANT: Ryan, Esther  
; APPLICANT: Lee, Seung-Wuk  
; TITLE OF INVENTION: BIOLOGICAL CONTROL OF NANOPARTICLES  
; FILE REFERENCE: 119927-1066  
; CURRENT APPLICATION NUMBER: US/10/254,446A  
; CURRENT FILING DATE: 2003-02-19  
; PRIOR APPLICATION NUMBER: 60/325,664  
; PRIOR FILING DATE: 2001-09-28  
; NUMBER OF SEQ ID NOS: 245  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 234  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: artificial sequence  
; FEATURE:  
; OTHER INFORMATION: peptide with peptide binding sequence retrieved from  
phage biopanning  
US-10-254-446A-234

Query Match 55.6%; Score 5; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRLHL 7  
 |||||  
Db 4 SRLHL 8

RESULT 4

US-09-852-910-200

; Sequence 200, Application US/09852910  
; Publication No. US20030096297A1  
; GENERAL INFORMATION:  
; APPLICANT: Hamm, Heidi  
; APPLICANT: Gilchrist, Annette  
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled  
Receptor Signaling  
; FILE REFERENCE: 2661-101  
; CURRENT APPLICATION NUMBER: US/09/852,910  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US 60/275,472  
; PRIOR FILING DATE: 2001-03-14  
; NUMBER OF SEQ ID NOS: 271  
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 200  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(11)  
; OTHER INFORMATION: Gs library peptide  
US-09-852-910-200

Query Match 55.6%; Score 5; DB 11; Length 11;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLHLR 8  
|||||  
Db 2 RLHLR 6

RESULT 5  
US-10-029-386-27913  
; Sequence 27913, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR GENE  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 27913  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC025745.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.8  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8  
US-10-029-386-27913

Query Match 55.6%; Score 5; DB 12; Length 18;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSRLH 6  
|||||  
Db 6 NSRLH 10

RESULT 6

US-09-962-756-677

```
; Sequence 677, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 677
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-962-756-677
```

```
Query Match          55.6%; Score 5; DB 12; Length 21;
Best Local Similarity 100.0%; Pred. No. 58;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      4 RLHLR 8
        |||||
Db      2 RLHLR 6
```

RESULT 7

US-09-563-222-39

```
; Sequence 39, Application US/09563222
; Publication No. US20030079253A1
; GENERAL INFORMATION:
; APPLICANT: Hiatt, Andrew
; APPLICANT: Hein, Mich B.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN
; TITLE OF INVENTION: EUKARYOTIC CELLS
; FILE REFERENCE: 310098.406
; CURRENT APPLICATION NUMBER: US/09/563,222
; CURRENT FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 197
```

; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-563-222-39

Query Match 44.4%; Score 4; DB 11; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRLH 6  
|||  
Db 3 SRLH 6

RESULT 8

US-10-052-578-211  
; Sequence 211, Application US/10052578  
; Publication No. US20030134787A1  
; GENERAL INFORMATION:  
; APPLICANT: Sloan-Kettering Institute for Cancer Research  
; APPLICANT: Rothman, James E.  
; APPLICANT: Mayhew, Mark  
; APPLICANT: Hoe, Mee H.  
; APPLICANT: Houghton, Alan  
; APPLICANT: Hartl, Ulrich  
; APPLICANT: Ouerfelli, Ouathak  
; APPLICANT: Moroi, Yoichi  
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES  
; FILE REFERENCE: 11746/46003  
; CURRENT APPLICATION NUMBER: US/10/052,578  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 08/961,707  
; PRIOR FILING DATE: 1997-10-31  
; NUMBER OF SEQ ID NOS: 321  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 211  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide in m13 coliphage  
US-10-052-578-211

Query Match 44.4%; Score 4; DB 12; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSRL 5  
|||  
Db 1 NSRL 4

RESULT 9

US-10-053-520-211  
; Sequence 211, Application US/10053520

```

; Publication No. US20030166530A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Hartl, Ulrich
; APPLICANT: Ouerfelli, Ouathék
; APPLICANT: Moroi, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46004
; CURRENT APPLICATION NUMBER: US/10/053,520
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 211
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide in m13 coliphage
US-10-053-520-211

```

```

Query Match          44.4%; Score 4; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      2 NSRL 5
        ||||
Db      1 NSRL 4

```

```

RESULT 10
US-10-053-498B-211
; Sequence 211, Application US/10053498B
; Publication No. US20030194409A1
; GENERAL INFORMATION:
; APPLICANT: Sloan-Kettering Institute for Cancer Research
; APPLICANT: Rothman, James E.
; APPLICANT: Mayhew, Mark
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Hartl, Ulrich
; APPLICANT: Ouerfelli, Ouathék
; APPLICANT: Moroi, Yoichi
; TITLE OF INVENTION: CONJUGATE HEAT SHOCK PROTEIN-BINDING PEPTIDES
; FILE REFERENCE: 11746/46002
; CURRENT APPLICATION NUMBER: US/10/053,498B
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 08/961,707
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 211

```

; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide in m13 coliphage  
US-10-053-498B-211

Query Match 44.4%; Score 4; DB 12; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSRL 5  
||||  
Db 1 NSRL 4

RESULT 11

US-10-140-555-9  
; Sequence 9, Application US/10140555  
; Publication No. US20020127227A1  
; GENERAL INFORMATION:  
; APPLICANT: Julie A. Abrahamson  
; APPLICANT: Stephen D. Holmes  
; APPLICANT: Jeffrey R. Jackson  
; TITLE OF INVENTION: RHAMM Antagonist Antibodies  
; FILE REFERENCE: P50857  
; CURRENT APPLICATION NUMBER: US/10/140,555  
; CURRENT FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: US/09/443,790  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/109,041  
; PRIOR FILING DATE: 1998-11-19  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)...(7)  
; OTHER INFORMATION: light chain CDR 2  
US-10-140-555-9

Query Match 44.4%; Score 4; DB 14; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRLH 6  
||||  
Db 3 SRLH 6

RESULT 12

US-10-139-496-11  
; Sequence 11, Application US/10139496  
; Publication No. US20030082646A1



```

; GENERAL INFORMATION:
; APPLICANT: Carey, Thomas E.
; APPLICANT: Nair, Thankum S.
; APPLICANT: Gray, Jennifer P.
; TITLE OF INVENTION: Antigenic Targets of Autoimmune Sensorineural Hearing
Loss (AISNHL) and
; TITLE OF INVENTION: Development of Tests for Diagnosis and Management of
AISNHL
; FILE REFERENCE: UM-6982
; CURRENT APPLICATION NUMBER: US/10/139,496
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 09/222,179
; PRIOR FILING DATE: 1998-12-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-139-496-11

```

```

Query Match          44.4%; Score 4; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      5 LHLR 8
        ||||
Db      4 LHLR 7

```

```

RESULT 13
US-09-832-723-110
; Sequence 110, Application US/09832723
; Patent No. US20020098524A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; APPLICANT: Chen, Yiyou
; APPLICANT: Murray, Christopher J.
; APPLICANT: Tijerina, Pilar
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING
; FILE REFERENCE: GC617-2
; CURRENT APPLICATION NUMBER: US/09/832,723
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,259
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 110
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptides screened from a phage display random
; OTHER INFORMATION: peptide library
US-09-832-723-110

```

Query Match 44.4%; Score 4; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 HLRC 9  
    ||||  
Db 6 HLRC 9

RESULT 14

US-09-952-432A-23

; Sequence 23, Application US/09952432A  
; Patent No. US20020150588A1  
; GENERAL INFORMATION:  
; APPLICANT: Allison, James P.  
; APPLICANT: Fasso, Marcella  
; APPLICANT: Shastri, Nilabh  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: SPAS-1 CANCER ANTIGEN  
; FILE REFERENCE: 018941-001110US  
; CURRENT APPLICATION NUMBER: US/09/952,432A  
; CURRENT FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: 60/234,472  
; PRIOR FILING DATE: 2000-09-21  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Predicted AA  
; OTHER INFORMATION: sequence SPAS-1 epitope from normal tissue

US-09-952-432A-23

Query Match 44.4%; Score 4; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 HLRC 9  
    ||||  
Db 6 HLRC 9

RESULT 15

US-10-303-331-110

; Sequence 110, Application US/10303331  
; Publication No. US20030152976A1  
; GENERAL INFORMATION:  
; APPLICANT: Janssen, Giselle G.  
; APPLICANT: Murray, Christopher J.  
; APPLICANT: Winetzky, Deborah S.  
; TITLE OF INVENTION: METHODS FOR SELECTIVE TARGETING  
; FILE REFERENCE: GC617-3  
; CURRENT APPLICATION NUMBER: US/10/303,331  
; CURRENT FILING DATE: 2002-11-25

; PRIOR APPLICATION NUMBER: US 09/832,723  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: US 60/197,259  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 110  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptides screened from a phage display random  
; OTHER INFORMATION: peptide library  
US-10-303-331-110

Query Match 44.4%; Score 4; DB 12; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 HLRC 9  
|||  
Db 6 HLRC 9

Search completed: November 13, 2003, 11:12:31  
Job time : 17.625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:26:01 ; Search time 8.8125 Seconds  
(without alignments)  
98.215 Million cell updates/sec

Title: US-09-228-866-1  
Perfect score: 9  
Sequence: 1 CNSRLHLRC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3752

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pirl:\*

2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3	33.3	7	2	PQ0663	membrane protein -
2	3	33.3	8	2	S59622	metallothionein is
3	3	33.3	10	2	PT0309	Ig heavy chain CRD
4	3	33.3	11	2	PT0217	T-cell receptor be
5	3	33.3	12	1	LFECPE	pyrE leader peptid
6	3	33.3	12	2	A29169	phospholipase A2 (
7	3	33.3	12	2	S65136	kallikrein K2 - hu
8	3	33.3	13	2	JH0460	corticostatic pept
9	3	33.3	13	2	B61620	locustamyotropin I
10	3	33.3	14	2	I54284	C1-inhibitor - hum
11	3	33.3	15	2	I49420	placental lactogen
12	3	33.3	15	2	S36891	ribosomal protein
13	3	33.3	15	2	PA0029	protein QA100012 -
14	3	33.3	15	2	G41299	T-cell receptor al
15	3	33.3	15	2	PH0775	T-cell receptor al
16	3	33.3	15	2	PH0779	T-cell receptor al
17	3	33.3	15	2	PH1455	T-cell receptor al
18	3	33.3	15	2	PH0770	T-cell receptor be
19	3	33.3	16	2	S35627	uvsX protein - pha
20	3	33.3	16	2	A28587	T-cell receptor be
21	3	33.3	17	2	I55226	myosin heavy chain
22	3	33.3	17	2	JP0068	ribosomal protein
23	3	33.3	17	2	S50742	proteinase inhibit
24	3	33.3	17	2	E28587	T-cell receptor be
25	3	33.3	17	2	A54205	heterogeneous ribo
26	3	33.3	17	2	H53284	T-cell receptor be
27	3	33.3	19	2	A48354	nonstructural prot
28	3	33.3	19	2	S63510	phosphonoacetaldeh
29	3	33.3	19	2	PC1324	hypothetical prote
30	3	33.3	19	2	PC1322	hypothetical prote
31	3	33.3	19	2	PH1353	Ig heavy chain DJ
32	3	33.3	20	2	C20554	hemocyanin subunit
33	3	33.3	20	2	H49034	nuclear antigen EB
34	3	33.3	20	2	I67551	monocyte chemotact
35	3	33.3	20	2	S56756	link protein - rat
36	3	33.3	21	2	I51224	somatotropin - chi
37	3	33.3	21	2	S08441	rev protein - huma
38	2	22.2	7	2	S19630	ribosomal protein
39	2	22.2	7	2	ECMUCR	catch-relaxing pep
40	2	22.2	7	2	A15398	choline oxidase (E
41	2	22.2	7	2	S42407	gramicidin S synth
42	2	22.2	7	2	PN0150	omega-gliadine 1'
43	2	22.2	7	2	A28340	myomodulin - Calif
44	2	22.2	7	2	PT0581	T-cell receptor be

## ALIGNMENTS

## RESULT 1

PQ0663

membrane protein - porcine epidemic diarrhea virus (isolate Belgian CV777)  
(fragment)

C;Species: porcine epidemic diarrhea virus

C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 08-Oct-1999

C;Accession: PQ0663

R;Bridgen, A.; Duarte, M.; Tobler, K.; Laude, H.; Ackermann, M.

J. Gen. Virol. 74, 1795-1804, 1993

A;Title: Sequence determination of the nucleocapsid protein gene of the porcine epidemic diarrhoea virus confirms that this virus is a coronavirus related to human coronavirus 229E and porcine transmissible gastroenteritis virus.

A;Reference number: JQ2191; MUID:93389433; PMID:8397280

A;Accession: PQ0663

A;Molecule type: mRNA

A;Residues: 1-7 &lt;BRI&gt;

A;Cross-references: GB:Z14976; NID:g311650; PIDN:CAA78699.1; PID:g584083

C;Comment: This virus is coronavirus related to human coronavirus 229E.

C;Keywords: membrane protein

Query Match 33.3%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LHL 7

|||

Db 4 LHL 6

## RESULT 2

S59622

metallothionein isoform a, cadmium-binding - Arianta arbustorum (terrestrial snail) (fragment)

C;Species: Arianta arbustorum

C;Date: 19-Mar-1997 #sequence\_revision 24-Oct-1997 #text\_change 07-May-1999

C;Accession: S59622

R;Berger, B.; Hunziker, P.E.; Hauer, C.R.; Birchler, N.; Dallinger, R.

Biochem. J. 311, 951-957, 1995

A;Title: Mass spectrometry and amino acid sequencing of two cadmium-binding metallothionein isoforms from the terrestrial gastropod Arianta arbustorum.

A;Reference number: S59621; MUID:96067616; PMID:7487956

A;Accession: S59622

A;Molecule type: protein

A;Residues: 1-8 &lt;BER&gt;

C;Superfamily: metallothionein

C;Keywords: chelation; metal binding; metal-thiolate cluster

Query Match 33.3%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNS 3  
|||  
Db 1 CNS 3

#### RESULT 3

PT0309

Ig heavy chain CRD3 region (clone 6-94) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C;Accession: PT0309

R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991

A;Title: Preferential utilization of specific immunoglobulin heavy chain  
diversity and joining segments in adult human peripheral blood B lymphocytes.

A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0309

A;Molecule type: DNA

A;Residues: 1-10 <YAM>

A;Experimental source: B lymphocyte

C;Keywords: heterotetramer; immunoglobulin

Query Match 33.3%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSR 4  
|||  
Db 5 NSR 7

#### RESULT 4

PT0217

T-cell receptor beta chain V-J region (4-1-E.2) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 30-May-1997

C;Accession: PT0217

R;Nakano, N.; Kikutani, H.; Nishimoto, H.; Kishimoto, T.  
J. Exp. Med. 173, 1091-1097, 1991

A;Title: T cell receptor V gene usage of islet beta cell-reactive T cells is not  
restricted in non-obese diabetic mice.

A;Reference number: PT0209; MUID:91217621; PMID:1902501

A;Accession: PT0217

A;Molecule type: mRNA

A;Residues: 1-11 <NAK>

C;Keywords: T-cell receptor

Query Match 33.3%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
|||  
Db 3 SRL 5

#### RESULT 5

LFECPPE

pyrE leader peptide - Escherichia coli

C;Species: Escherichia coli

C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 31-Mar-1993

C;Accession: A30400; A05110; Q00495

R;Poulsen, P.; Bonekamp, F.; Jensen, K.F.

EMBO J. 3, 1783-1790, 1984

A;Title: Structure of the Escherichia coli pyrE operon and control of pyrE expression by a UTP modulated intercistronic attenuation.

A;Reference number: A30400; MUID:85003588; PMID:6207018

A;Accession: A30400

A;Molecule type: DNA

A;Residues: 1-12 <POU1>

R;Poulsen, P.; Jensen, K.F.; Valentin-Hansen, P.; Carlsson, P.; Lundberg, L.G.

Eur. J. Biochem. 135, 223-229, 1983

A;Title: Nucleotide sequence of the Escherichia coli pyrE gene and of the DNA in front of the protein-coding region.

A;Reference number: A05110; MUID:83287414; PMID:6349999

A;Accession: A05110

A;Molecule type: DNA

A;Residues: 1-12 <POU2>

C;Genetics:

A;Gene: pyrE-LP

A;Map position: 82 min

C;Superfamily: pyrE leader peptide

Query Match 33.3%; Score 3; DB 1; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.8e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SRL 5

|||

Db 2 SRL 4

RESULT 6

A29169

phospholipase A2 (EC 3.1.1.4) precursor - sheep (fragment)

C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C;Date: 02-Jun-1988 #sequence\_revision 02-Jun-1988 #text\_change 31-Oct-1997

C;Accession: A29169

R;Dutilh, C.E.; Van Doren, P.J.; Verheul, F.E.A.M.; De Haas, G.H.

Eur. J. Biochem. 53, 91-97, 1975

A;Title: Isolation and properties of prophospholipase A2 from ox and sheep pancreas.

A;Reference number: A94661

A;Accession: A29169

A;Molecule type: protein

A;Residues: 1-12 <DUT>

C;Superfamily: phospholipase A2

C;Keywords: carboxylic ester hydrolase; pyroglutamic acid

F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 33.3%; Score 3; DB 2; Length 12;

Best Local Similarity 100.0%; Pred. No. 1.8e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            2 NSR 4  
             |||  
Db            5 NSR 7

RESULT 7

S65136

kallikrein K2 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 07-May-1999

C;Accession: S65136

R;Depertes, D.; Chapdelaine, P.; Tremblay, R.R.; Brunet, C.; Berton, J.;  
Hebert, J.; Lazure, C.; Dube, J.Y.

Biochim. Biophys. Acta 1245, 311-316, 1995

A;Title: Isolation of prostatic kallikrein hK2, also known as hGK-1, in human  
seminal plasma.

A;Reference number: S65136; MUID:96125726; PMID:8541306

A;Accession: S65136

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-12 <DEP>

Query Match                    33.3%;   Score 3;   DB 2;   Length 12;  
Best Local Similarity       100.0%;   Pred. No. 1.8e+03;  
Matches       3;   Conservative       0;   Mismatches       0;   Indels       0;   Gaps       0;

Qy            5 LHL 7  
             |||  
Db            7 LHL 9

RESULT 8

JH0460

corticostatic peptide GP-CS3 - guinea pig

C;Species: Cavia porcellus (guinea pig)

C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 18-Aug-2000

C;Accession: JH0460

R;Hu, J.; Bennett, H.P.J.; Lazure, C.; Solomon, S.

Biochem. Biophys. Res. Commun. 180, 558-565, 1991

A;Title: Isolation and characterization of corticostatic peptides from guinea  
pig bone marrow.

A;Reference number: JH0458; MUID:92062075; PMID:1659400

A;Accession: JH0460

A;Molecule type: protein

A;Residues: 1-13 <HUJ>

A;Experimental source: bone marrow

A;Note: this is a dimer having an antiparallel configuration

C;Comment: This peptide belongs to a family of Cys-rich, cationic peptides of  
low molecular weight.

C;Comment: This peptide has antimicrobial activity by a non-oxygen-dependent  
mechanism.

C;Superfamily: unassigned animal peptides

F;5/Disulfide bonds: interchain (to 13) #status experimental

F;7/Disulfide bonds: interchain (to 11) #status experimental

F;11/Disulfide bonds: interchain (to 7) #status experimental

F;13/Disulfide bonds: interchain (to 5) #status experimental



Query Match 33.3%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RLH 6  
|||  
Db 8 RLH 10

RESULT 9

B61620

locustamyotropin IV - migratory locust

C;Species: Locusta migratoria (migratory locust)

C;Date: 21-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 11-Jul-1997

C;Accession: B61620

R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; Kochansky, J.P.; De Loof, A.

Insect Biochem. Mol. Biol. 22, 447-452, 1992

A;Title: Isolation, identification and synthesis of locustamyotropin III and IV, two additional neuropeptides of Locusta migratoria: members of the locustamyotropin peptide family.

A;Reference number: A61620

A;Accession: B61620

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-13 <SCH>

C;Keywords: amidated carboxyl end; neuropeptide

F;13/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 33.3%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RLH 6  
|||  
Db 1 RLH 3

RESULT 10

I54284

C1-inhibitor - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 21-Jul-2000

C;Accession: I54284

R;Siddique, Z.; McPhaden, A.R.; McCluskey, D.; Whaley, K.

Hum. Hered. 42, 231-234, 1992

A;Title: A single base deletion from the C1-inhibitor gene causes type I hereditary angio-oedema.

A;Reference number: I54284; MUID:92380682; PMID:1339401

A;Accession: I54284

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-14 <RES>

A;Cross-references: GB:S44615; NID:g254386; PIDN:AAB23055.1; PID:g254387

Query Match 33.3%; Score 3; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRC 9  
|||  
Db 5 LRC 7

RESULT 11

I49420

placental lactogen I - western wild mouse (fragment)

C;Species: Mus spretus (western wild mouse)

C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 11-Jan-2000

C;Accession: I49420

R;Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.

Mamm. Genome 5, 349-355, 1994

A;Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.

A;Reference number: I48934; MUID:94319082; PMID:8043949

A;Accession: I49420

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-15 <RES>

A;Cross-references: EMBL:U05735; NID:g497071; PIDN:AAB60476.1; PID:g497072

C;Superfamily: prolactin

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRC 9  
|||  
Db 5 LRC 7

RESULT 12

S36891

ribosomal protein - Mycobacterium bovis (fragment)

C;Species: Mycobacterium bovis

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 13-Jan-1995

C;Accession: S36891

R;Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.

FEBS Lett. 331, 9-14, 1993

A;Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobacterium bovis BCG.

A;Reference number: S36887; MUID:94009653; PMID:8405418

A;Accession: S36891

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-15 <OHA>

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSR 4  
|||  
Db 8 NSR 10

# RESULT 13

PA0029

protein QA100012 - Arabidopsis thaliana (fragment)

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 30-Jun-1992 #sequence\_revision 06-Jan-1995 #text\_change 06-Jun-1997

C;Accession: PA0029

R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.

submitted to JIPID, July 1994

A;Description: Separation and characterization of Arabidopsis proteins by two-dimensional gel electrophoresis.

A;Reference number: PA0001

A;Accession: PA0029

A;Molecule type: protein

A;Residues: 1-15 <KAM>

A;Experimental source: callus

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
|||  
Db 4 SRL 6

# RESULT 14

G41299

T-cell receptor alpha chain precursor J region (39) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 05-Nov-1999

C;Accession: G41299

R;Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.;

Panayi, G.; Steinmetz, M.

Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991

A;Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumatoid arthritis is polyclonal.

A;Reference number: A41299; MUID:92020887; PMID:1656449

A;Accession: G41299

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-15 <UEM>

A;Cross-references: GB:S57457; NID:g236330; PIDN:AAB19962.1; PID:g236331

C;Keywords: T-cell receptor

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
|||  
Db 9 SRL 11

# RESULT 15

PH0775

T-cell receptor alpha chain (B28) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997  
C;Accession: PH0775  
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
J. Exp. Med. 174, 1371-1383, 1991  
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.  
A;Reference number: PH0746; MUID:92078846; PMID:1836010  
A;Accession: PH0775  
A;Molecule type: mRNA  
A;Residues: 1-15 <CAS>  
A;Cross-references: EMBL:X60871  
A;Experimental source: T lymphocyte  
C;Keywords: T-cell receptor

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLH 6  
|||  
Db 11 RLH 13

Search completed: November 13, 2003, 10:39:52  
Job time : 9.8125 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:55:06 ; Search time 4.875 Seconds  
(without alignments)  
86.819 Million cell updates/sec

Title: US-09-228-866-1  
Perfect score: 9  
Sequence: 1 CNSRLHLRC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1246

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3	33.3	12	1	YZPY_ECOLI	P17776 escherichia
2	3	33.3	13	1	LMT4_LOCFI	P41490 locusta mig
3	3	33.3	16	1	UVSX_BPT6	Q06728 bacterioph
4	3	33.3	21	1	FIBB_CEREL	P14468 cervus elap
5	3	33.3	21	1	FIBB_MUNMU	P14475 muntiacus m
6	3	33.3	21	1	FIBB_ODOHE	P14476 odocoileus
7	3	33.3	21	1	FIBB_RANTA	P14479 rangifer ta
8	3	33.3	21	1	REV_HV2D2	P15830 human immun
9	2	22.2	7	1	CARP_MYTED	P10420 mytilus edu
10	2	22.2	7	1	CHOX_ALCSP	P16101 alcaligenes
11	2	22.2	7	1	FAR1_HELTI	P41871 helisoma tr
12	2	22.2	7	1	FAR1_MACRS	P83274 macrobrachi
13	2	22.2	7	1	FAR1_PROCL	P38499 procambarus
14	2	22.2	7	1	FAR2_ASCSU	P31890 ascaris suu
15	2	22.2	7	1	FAR2_PROCL	P38498 procambarus
16	2	22.2	8	1	ALL3_CYDPO	P82154 cydia pomon
17	2	22.2	8	1	FAR1_PANRE	P41872 panagrellus
18	2	22.2	8	1	FAR1_PENMO	P83316 penaeus mon
19	2	22.2	8	1	FAR2_MACRS	P83275 macrobrachi
20	2	22.2	8	1	FAR3_HOMAM	P41486 homarus ame
21	2	22.2	8	1	FAR4_HOMAM	P41487 homarus ame
22	2	22.2	8	1	FAR4_MACRS	P83277 macrobrachi
23	2	22.2	8	1	LCK1_LEUMA	P21140 leucophaea
24	2	22.2	8	1	LCK3_LEUMA	P21142 leucophaea
25	2	22.2	8	1	LMT2_LOCFI	P22396 locusta mig
26	2	22.2	8	1	LPK_LEUMA	P13049 leucophaea
27	2	22.2	8	1	LPMS_STAEP	P23211 staphylococ
28	2	22.2	8	1	PPK2_PERAM	P82692 periplaneta
29	2	22.2	8	1	PPK3_PERAM	P82618 periplaneta
30	2	22.2	9	1	CCAP_CARMA	P38556 carcinus ma
31	2	22.2	9	1	FAR2_PANRE	P41873 panagrellus
32	2	22.2	9	1	FAR3_MACRS	P83276 macrobrachi
33	2	22.2	9	1	FAR3_PENMO	P83318 penaeus mon
34	2	22.2	9	1	FAR4_PENMO	P83319 penaeus mon
35	2	22.2	9	1	FAR5_PENMO	P83320 penaeus mon
36	2	22.2	9	1	FAR6_MACRS	P83279 macrobrachi
37	2	22.2	9	1	FAR8_MACRS	P83281 macrobrachi
38	2	22.2	9	1	FAR9_ASCSU	P43172 ascaris suu
39	2	22.2	9	1	FARP_CALSI	P38495 callinectes
40	2	22.2	9	1	LMT3_LOCFI	P41489 locusta mig
41	2	22.2	9	1	MOSH_CLYJA	P19852 clypeaster
42	2	22.2	9	1	PPK1_PERAM	P82691 periplaneta
43	2	22.2	9	1	PTSP_BOMMO	P82003 bombyx mori
44	2	22.2	9	1	TAL3_PICJA	P17441 pichia jadi
45	2	22.2	10	1	AKHX_LOCFI	P81626 locusta mig

# ALIGNMENTS

## RESULT 1

YZPY\_ECOLI

ID YZPY\_ECOLI STANDARD; PRT; 12 AA.  
AC P17776;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Hypothetical pyrE leader peptide.  
GN PYRL OR PYRE-LP.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85003588; PubMed=6207018;  
RA Poulsen P., Bonekamp F., Jensen K.F.;  
RT "Structure of the Escherichia coli pyrE operon and control of pyrE  
RT expression by a UTP modulated intercistronic attenuation.";  
RL EMBO J. 3:1783-1790(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=83287414; PubMed=6349999;  
RA Poulsen P., Jensen K.F., Valentin-Hansen P., Carlsson P.,  
RA Lundberg L.G.;  
RT "Nucleotide sequence of the Escherichia coli pyrE gene and of the DNA  
RT in front of the protein-coding region.";  
RL Eur. J. Biochem. 135:223-229(1983).  
CC -!- CAUTION: THIS SEQUENCE, ACCORDING TO THE ECOSEQ DATABASE (K. RUDD)  
CC IS PROBABLY NOT A REAL PROTEIN; THEREFORE THIS ENTRY WILL PROBABLY  
CC BE DELETED IN FUTURE RELEASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X00781; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; V01578; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A30400; LFECPE.  
KW Hypothetical protein.  
SQ SEQUENCE 12 AA; 1542 MW; C4291FA437A2C9C9 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5

Db

|||  
2 SRL 4

RESULT 2

LMT4\_LOCMI

ID LMT4\_LOCMI STANDARD; PRT; 13 AA.  
AC P41490;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Locustamyotropin 4 (LOM-MT-4).  
OS Locusta migratoria (Migratory locust).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
OX NCBI\_TaxID=7004;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC TISSUE=Brain;  
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,  
RA de Loof A.;  
RT "Isolation, identification and synthesis of locustamyotropin III and  
RT IV, two additional neuropeptides of Locusta migratoria: members of the  
RT locustamyotropin peptide family.";  
RL Insect Biochem. Mol. Biol. 22:447-452(1992).  
CC -!- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTILE ACTIVITY  
CC (MYOTROPIC ACTIVITY). LOM-MT IV SEEMS TO BE A MORE POTENT MUSCLE  
CC STIMULATOR THAN LOM-MT I, II AND III.  
CC -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.  
DR PIR; B61620; B61620.  
DR InterPro; IPR001484; Pyrokinin.  
DR PROSITE; PS00539; PYROKININ; 1.  
KW Neuropeptide; Amidation; Pyrokinin.  
FT MOD\_RES 13 13 AMIDATION.  
SQ SEQUENCE 13 AA; 1553 MW; 20861943824D6698 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLH 6

|||

Db 1 RLH 3

RESULT 3

UVSX\_BPT6

ID UVSX\_BPT6 STANDARD; PRT; 16 AA.  
AC Q06728;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Recombination and repair protein (Fragment).  
GN UVSX.  
OS Bacteriophage T6.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;

OC T4-like viruses.  
 OX NCBI\_TaxID=10666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93219141; PubMed=8464751;  
 RA Winkler M., Rueger W.;  
 RT "Cloning and sequencing of the genes of  
 RT beta-glucosyl-HMC-alpha-glucosyl-transferases of bacteriophages T2  
 RT and T6.";  
 RL Nucleic Acids Res. 21:1500-1500(1993).  
 CC -!- FUNCTION: IMPORTANT IN GENETIC RECOMBINATION, DNA REPAIR, AND  
 CC REPLICATION. POSSESSES PAIRING AND STRAND-TRANSFER ACTIVITY.  
 CC INTERACTS WITH DDA AND GENE 32 PROTEINS.  
 CC -!- SIMILARITY: Belongs to the recA family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X68725; CAA48668.1; -.  
 DR PIR; S35627; S35627.  
 KW DNA damage; DNA replication; DNA recombination; DNA repair;  
 KW ATP-binding.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1721 MW; 48164C95D76F3CB4 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
 |||  
 Db 8 SRL 10

#### RESULT 4

##### FIBB\_CEREL

ID FIBB\_CEREL STANDARD; PRT; 21 AA.  
 AC P14468;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Cervus elaphus (Red deer), and  
 OS Cervus elaphus nelsoni (American elk).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
 OC Cervidae; Cervinae; Cervus.  
 OX NCBI\_TaxID=9860, 9864;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=C.elaphus;



RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals.";  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=C.e.nelsoni;  
 RA Mross G.A., Doolittle R.F.;  
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
 RL Arch. Biochem. Biophys. 122:674-684(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation; Pyrrolidone carboxylic acid.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2558 MW; FCEE745D98931627 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LHL 7  
 |||  
 Db 16 LHL 18

RESULT 5  
 FIBB\_MUNMU  
 ID FIBB\_MUNMU STANDARD; PRT; 21 AA.  
 AC P14475;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Muntiacus muntjak (Muntjak).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
 OC Cervidae; Muntiacinae; Muntiacus.  
 OX NCBI\_TaxID=9888;  
 RN [1]  
 RP SEQUENCE.  
 RA Mross G.A., Doolittle R.F.;  
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
 RL Arch. Biochem. Biophys. 122:674-684(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET

CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation; Pyrrolidone carboxylic acid.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2514 MW; FCEE75188F0C1627 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LHL 7  
 |||  
 Db 16 LHL 18

#### RESULT 6

##### FIBB\_ODOHE

ID FIBB\_ODOHE STANDARD; PRT; 21 AA.  
 AC P14476;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Odocoileus hemionus (Mule deer) (Black-tailed deer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;  
 OC Cervidae; Odocoileinae; Odocoileus.  
 OX NCBI\_TaxID=9872;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=67209145; PubMed=6033721;  
 RA Doolittle R.F., Schubert D., Schwartz S.A.;  
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.  
 RT Dromedary camel, mule deer, and cape buffalo."  
 RL Arch. Biochem. Biophys. 118:456-467(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.

KW Blood coagulation; Plasma; Sulfation; Pyrrolidone carboxylic acid.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2496 MW; FCF562C51A0C1627 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LHL 7  
 |||  
 Db 16 LHL 18

# RESULT 7

FIBB\_RANTA  
 ID FIBB\_RANTA STANDARD; PRT; 21 AA.  
 AC P14479;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Rangifer tarandus (Reindeer) (Caribou).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidea;  
 OC Cervidae; Odocoileinae; Rangifer.  
 OX NCBI\_TaxID=9870;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals."  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation; Pyrrolidone carboxylic acid.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2510 MW; FCF562C45F0C1627 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            5 LHL 7  
               |||  
 Db           16 LHL 18

RESULT 8

REV\_HV2D2

ID REV\_HV2D2            STANDARD;            PRT;            21 AA.  
 AC P15830;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE REV protein (Anti-repression transactivator protein) (ART/TRS)  
 DE (Fragment).  
 GN REV.  
 OS Human immunodeficiency virus type 2 (isolate D205,7) (HIV-2).  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11716;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90081881; PubMed=2594088;  
 RA Dietrich U., Adamski M., Kreutz R., Seipp A., Kuehnel H.,  
 RA Ruebsamen-Waigmann H.;  
 RT "A highly divergent HIV-2-related isolate."  
 RL Nature 342:948-950(1989).  
 CC -!- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE  
 CC        NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.  
 CC -!- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.  
 CC -!- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED  
 CC        BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X61240; -, NOT\_ANNOTATED\_CDS.  
 DR PIR; S08441; S08441.  
 DR HIV; X16109; REV\$2D205.  
 DR InterPro; IPR000625; REV\_protein.  
 DR Pfam; PF00424; REV; 1.  
 KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.  
 FT NON\_TER            21            21  
 SQ SEQUENCE    21 AA;    2503 MW;    E620E225CC5BFF24 CRC64;

Query Match            33.3%;    Score 3;    DB 1;    Length 21;  
 Best Local Similarity    100.0%;    Pred. No. 8.8e+02;  
 Matches        3;    Conservative        0;    Mismatches        0;    Indels        0;    Gaps        0;

Qy            5 LHL 7  
               |||  
 Db           15 LHL 17

# RESULT 9

## CARP\_MYTED

ID CARP\_MYTED STANDARD; PRT; 7 AA.  
AC P10420;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-MAR-1989 (Rel. 10, Last annotation update)  
DE Catch-relaxing peptide (CARP).  
OS Mytilus edulis (Blue mussel).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;  
OC Mytiloidea; Mytilidae; Mytilus.  
OX NCBI\_TaxID=6550;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=88052022; PubMed=3676797;  
RA Hirata T., Kubota I., Takabatake I., Kawahara A., Shimamoto N.,  
RA Muneoka Y.;  
RT "Catch-relaxing peptide isolated from Mytilus pedal ganglia.";  
RL Brain Res. 422:374-376(1987).  
CC -!- FUNCTION: THIS PEPTIDE EXHIBITS BOTH POTENTIATING (CONTRACTION)  
CC AND INHIBITORY (RELAXATION) EFFECTS ON THE ANTERIOR BYSSUS  
CC RETRACTOR MUSCLE.  
DR PIR; A29342; ECMUCR.  
KW Hormone; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 831 MW; 6734072687669DB0 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RL 5  
||  
Db 6 RL 7

# RESULT 10

## CHOX\_ALCSP

ID CHOX\_ALCSP STANDARD; PRT; 7 AA.  
AC P16101;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DE Choline oxidase (EC 1.1.3.17) (Fragment).  
OS Alcaligenes sp.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Alcaligenaceae; Alcaligenes.  
OX NCBI\_TaxID=512;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=81006769; PubMed=6997283;  
RA Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;  
RT "Identification and properties of the prosthetic group of choline  
RT oxidase from Alcaligenes sp.";  
RL J. Biochem. 88:197-203(1980).  
CC -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).  
DR PIR; A15398; A15398.

KW Oxidoreductase.  
FT NON\_TER 7 7  
SQ SEQUENCE 7 AA; 839 MW; 7415B1E457644AC0 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SR 4  
||  
Db 6 SR 7

RESULT 11

FAR1\_HELTI

ID FAR1\_HELTI STANDARD; PRT; 7 AA.  
AC P41871;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FMRFamide-like neuropeptide GDPFLRF-amide.  
OS Helisoma trivolvis (Snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
OC Lymnaeoidae; Planorbidae; Helisoma.  
OX NCBI\_TaxID=27815;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=94286417; PubMed=7912428;  
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
RT trivolvis.";  
RL Peptides 15:31-36(1994).  
CC -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING  
CC THE KIDNEY, MANTLE AND SKIN.  
CC -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LR 8  
||  
Db 5 LR 6

RESULT 12

FAR1\_MACRS

ID FAR1\_MACRS STANDARD; PRT; 7 AA.  
AC P83274;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE FMRFamide-like neuropeptide FLP1 (DRNFLRF-amide).  
 OS Macrobrachium rosenbergii (Giant fresh water prawn).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
 OC Palaemonoidea; Palaemonidae; Macrobrachium.  
 OX NCBI\_TaxID=79674;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Eyestalk;  
 RA Sithigorngul P., Saraithongkum W., Jaideechoey S., Longyant S.,  
 RA Sithigorngul W.;  
 RT "Novel FMRFamide-like neuropeptides from the eyestalk of the giant  
 RT freshwater prawn Macrobrachium rosenbergii.";  
 RL Comp. Biochem. Physiol. 120B:587-595(1998).  
 CC -!- MASS SPECTROMETRY: MW=965.7; METHOD=MALDI.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR GO; GO:0007218; P:neuropeptide signaling pathway; TAS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 7 7 AMIDATION.  
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LR 8  
 ||  
 Db 5 LR 6

#### RESULT 13

##### FAR1\_PROCL

ID FAR1\_PROCL STANDARD; PRT; 7 AA.  
 AC P38499;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Cardioexcitatory FMRFamide homolog NF1.  
 OS Procambarus clarkii (Red swamp crayfish).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;  
 OC Astacoidea; Cambaridae; Procambarus.  
 OX NCBI\_TaxID=6728;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Pericardial organs;  
 RX MEDLINE=93248032; PubMed=8387183;  
 RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;  
 RT "Isolation of two FMRFamide-related peptides from crayfish  
 RT pericardial organs.";  
 RL Peptides 14:137-143(1993).  
 CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS  
 CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF  
 CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)

CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 966 MW; 69D40729C4540420 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LR 8  
||  
Db 5 LR 6

RESULT 14

FAR2\_ASCSU

ID FAR2\_ASCSU STANDARD; PRT; 7 AA.  
AC P31890;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE FMRFamide-like neuropeptide AF2.  
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides), and  
OS Panagrellus redivivus.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;  
OC Ascarididae; Ascaris.  
OX NCBI\_TaxID=6253, 6233;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=A.suum;  
RX MEDLINE=93324431; PubMed=8332542;  
RA Cowden C., Stretton A.O.W.;  
RT "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";  
RL Peptides 14:423-430(1993).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=P.redivivus;  
RX MEDLINE=95060998; PubMed=7970891;  
RA Maule A.G., Shaw C., Bowman J.W.;  
RT "The FMRFamide-like neuropeptide AF2 (Ascaris suum) is present in the  
RT free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";  
RL Parasitology 109:351-356(1994).  
CC -!- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.  
CC -!- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF  
CC GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 992 MW; 69D4073B5B11E350 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LR 8  
||



Db 5 LR 6

RESULT 15

FAR2\_PROCL

ID FAR2\_PROCL STANDARD; PRT; 7 AA.  
AC P38498;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Cardioexcitatory FMRFamide homolog DF2.  
OS Procambarus clarkii (Red swamp crayfish).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;  
OC Astacoidea; Cambaridae; Procambarus.  
OX NCBI\_TaxID=6728;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Pericardial organs;  
RX MEDLINE=93248032; PubMed=8387183;  
RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;  
RT "Isolation of two FMRFamide-related peptides from crayfish  
pericardial organs.";  
RL Peptides 14:137-143(1993).  
CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS  
CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF  
CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LR 8  
||  
Db 5 LR 6

Search completed: November 13, 2003, 10:33:55  
Job time : 4.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:58:36 ; Search time 22.125 Seconds  
(without alignments)  
104.971 Million cell updates/sec

Title: US-09-228-866-1  
Perfect score: 9

Sequence: 1 CNSRLHLRC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4	44.4	20	11	Q8R4P6	Q8r4p6 mus musculu
2	3	33.3	9	4	Q9UC36	Q9uc36 homo sapien
3	3	33.3	9	11	Q61723	Q61723 mus musculu
4	3	33.3	10	4	Q9UN90	Q9un90 homo sapien
5	3	33.3	10	15	Q8UT83	Q8ut83 human immun
6	3	33.3	11	2	P77404	P77404 escherichia
7	3	33.3	12	15	Q85631	Q85631 avian carci
8	3	33.3	13	6	Q9TRW6	Q9trw6 bos taurus
9	3	33.3	13	11	Q9QVI3	Q9qvi3 cavia (guin
10	3	33.3	13	15	Q85645	Q85645 mouse mamma
11	3	33.3	14	2	P83159	P83159 anabaena sp
12	3	33.3	14	4	Q16045	Q16045 homo sapien
13	3	33.3	14	4	Q8TD29	Q8td29 homo sapien

14	3	33.3	14	11	Q9QV81	Q9qv81 rattus sp.
15	3	33.3	15	2	Q9R545	Q9r545 mycobacteri
16	3	33.3	15	4	Q9Y4Z9	Q9y4z9 homo sapien
17	3	33.3	15	4	P78533	P78533 homo sapien
18	3	33.3	15	8	Q95751	Q95751 brachylophu
19	3	33.3	15	8	Q95771	Q95771 ctenosaura
20	3	33.3	15	8	Q95952	Q95952 sauromalus
21	3	33.3	15	8	Q95773	Q95773 conolophus
22	3	33.3	15	8	Q95949	Q95949 sauromalus
23	3	33.3	15	8	Q95770	Q95770 cyclura ric
24	3	33.3	15	11	Q62544	Q62544 mus spretus
25	3	33.3	16	4	Q9UP51	Q9up51 homo sapien
26	3	33.3	16	4	Q96RT5	Q96rt5 homo sapien
27	3	33.3	16	8	Q8SL50	Q8sl50 aeonium mas
28	3	33.3	17	2	O34216	O34216 sphingomona
29	3	33.3	17	4	Q9HB76	Q9hb76 homo sapien
30	3	33.3	17	10	Q9S8K1	Q9s8k1 solanum tub
31	3	33.3	17	10	Q94FC0	Q94fc0 arabidopsis
32	3	33.3	17	11	Q61932	Q61932 mus musculu
33	3	33.3	17	12	Q64973	Q64973 alfalfa mos
34	3	33.3	17	12	Q64974	Q64974 alfalfa mos
35	3	33.3	18	4	Q8NED1	Q8ned1 homo sapien
36	3	33.3	18	6	O46592	O46592 capra hircu
37	3	33.3	19	2	Q53545	Q53545 shigella so
38	3	33.3	19	2	Q9R4G3	Q9r4g3 pseudomonas
39	3	33.3	19	4	O95578	O95578 homo sapien
40	3	33.3	19	4	Q96ER8	Q96er8 homo sapien
41	3	33.3	19	5	Q8T8B6	Q8t8b6 ciona intes
42	3	33.3	19	6	O46472	O46472 felis silve
43	3	33.3	19	6	Q95J98	Q95j98 sus scrofa
44	3	33.3	19	8	Q9G466	Q9g466 brassica ju
45	3	33.3	19	10	Q8GS32	Q8gs32 hordeum vul

# ALIGNMENTS

## RESULT 1

Q8R4P6

ID Q8R4P6 PRELIMINARY; PRT; 20 AA.

AC Q8R4P6;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Fus 2 (Fragment).

GN FUS2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C129/Sv;

RA Shuttleworth T.L., Wilson M.D., Wicklow B.A., Wilkins J.,

RA Triggs-Raine B.L.;

RT "Characterization of the Murine Hyaluronidase Gene Region Reveals

RT Complex Organization and Co-transcription of Hyal1 with Downstream

RT Genes, Fus2 and Hyal3.";  
 RL J. Biol. Chem. 0:0-0(2002).  
 DR EMBL; AF417496; AAM14429.1; -.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2333 MW; ADC481272C579A59 CRC64;

Query Match 44.4%; Score 4; DB 11; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLHL 7  
 ||||  
 Db 15 RLHL 18

# RESULT 2

Q9UC36

ID Q9UC36 PRELIMINARY; PRT; 9 AA.  
 AC Q9UC36;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE 28 kDa heat shock protein homolog fragment 1 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92218434; PubMed=1560006;  
 RA Kato K., Shinohara H., Goto S., Inaguma Y., Morishita R., Asano T.;  
 RT "Copurification of small heat shock protein with alpha B crystallin  
 RT from human skeletal muscle.";  
 RL J. Biol. Chem. 267:7718-7725(1992).  
 FT NON\_TER 1 1  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1220 MW; 26933415B1F77B43 CRC64;

Query Match 33.3%; Score 3; DB 4; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
 |||  
 Db 6 SRL 8

# RESULT 3

Q61723

ID Q61723 PRELIMINARY; PRT; 9 AA.  
 AC Q61723;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE NF-kappa-B DNA-binding subunit (Fragment).  
 GN NFKB1 OR NF-KAPPA-B.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/cByJ; TISSUE=Spleen;  
 RX MEDLINE=90367113; PubMed=2203532;  
 RA Ghosh S., Gifford A.M., Riviere L.R., Tempst P., Nolan G.P.,  
 RA Baltimore D.;  
 RT "Cloning of the p50 DNA binding subunit of NF-kappa-B: Homology to rel  
 RT and dorsal.";  
 RL Cell 62:1019-1029(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/cByJ; TISSUE=Spleen;  
 RX MEDLINE=94156215; PubMed=8112620;  
 RA Huo L., Chung W.H., Rothstein T.L.;  
 RT "C-terminal sequence of the NF-kappa-B p50 precursor from primary  
 RT murine B-lymphocytes.";  
 RL Gene 139:287-288(1994).  
 DR EMBL; L13466; AAC37644.1; -.  
 DR MGD; MGI:97312; Nfkb1.  
 KW DNA-binding.  
 FT NON\_TER 1 1  
 FT CONFLICT 5 5 A -> P (IN REF. 1).  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 925 MW; 300821E72DC1B408 CRC64;

Query Match 33.3%; Score 3; DB 11; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LHL 7  
 |||  
 Db 6 LHL 8

#### RESULT 4

Q9UN90

ID Q9UN90 PRELIMINARY; PRT; 10 AA.  
 AC Q9UN90;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE Canalicular multispecific organic anion transporter (Fragment).  
 GN CMOAT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tanaka T., Uchiumi T., Hinoshita E., Inokuchi A., Toh S., Wada M.,  
 RA Nomoto M., Kohno K., Kuwano M.;  
 RT "Sequence analysis and functional characterization of the 5'-flanking  
 RT region of the human canalicular multispecific organic anion  
 RT transporter/multidrug resistance protein 2 (cMOAT/MRP2) gene.";

RL Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF144630; AAD47599.1; -.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1219 MW; 76F28CB44EB9C33B CRC64;

Query Match 33.3%; Score 3; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNS 3  
|||  
Db 6 CNS 8

RESULT 5

Q8UT83

ID Q8UT83 PRELIMINARY; PRT; 10 AA.  
AC Q8UT83;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Vpu protein.  
GN VPU.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=00BW1795.6;  
RA Novitsky V.A., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,  
RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,  
RA Foley B.T., Gaolekwe S., Rybak N., Gaseitsiwe S., Vannberg F.,  
RA Marlink R., Lee T.-H., Essex M.;  
RT "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS  
RT vaccine design.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF443097; AAL34766.1; -.  
SQ SEQUENCE 10 AA; 1264 MW; 91E52CB33321B37A CRC64;

Query Match 33.3%; Score 3; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
|||  
Db 8 SRL 10

RESULT 6

P77404

ID P77404 PRELIMINARY; PRT; 11 AA.  
AC P77404;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE DNA sequence downstream of the ECOPRRI HSD locus (Fragment).  
GN HSDR.

OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97206151; PubMed=9157244;  
 RA Tyndall C., Lehnherr H., Sandmeier U., Kulik E., Bickle T.A.;  
 RT "The type IC hsd loci of the enterobacteria are flanked by DNA with  
 RT high homology to the phage P1 genome: implications for the evolution  
 RT and spread of DNA restriction systems.";  
 RL Mol. Microbiol. 23:729-736(1997).  
 DR EMBL; X98145; CAA66840.1; -.  
 DR EMBL; X98144; CAA66839.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 11 AA; 1259 MW; 714AB092A4072734 CRC64;

Query Match 33.3%; Score 3; DB 2; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
 |||  
 Db 3 SRL 5

# RESULT 7

Q85631

ID Q85631 PRELIMINARY; PRT; 12 AA.  
 AC Q85631;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE MH2, proviral DNA, myc to 3' LTR (Fragment).  
 OS Avian carcinoma virus.  
 OC Viruses; Retroviroidea; Retroviridae; Alpharetrovirus.  
 OX NCBI\_TaxID=11958;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85033920; PubMed=6092695;  
 RA Sutcliffe P., Jansen H.W., Bister K., Rapp U.R.;  
 RT "3'-terminal region of avian carcinoma virus MH2 shares sequence  
 RT elements with avian sarcoma viruses Y73 and SR-A.";  
 RL J. Virol. 52:703-705(1984).  
 DR EMBL; K03100; AAA42388.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 12 AA; 1466 MW; 72E4B884F30736DB CRC64;

Query Match 33.3%; Score 3; DB 15; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSR 4  
 |||  
 Db 9 NSR 11

# RESULT 8

Q9TRW6

ID Q9TRW6 PRELIMINARY; PRT; 13 AA.  
AC Q9TRW6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE 25 kDa protein P25, peptide F3 (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=91372400; PubMed=1909972;  
RA Takahashi M., Tomizawa K., Ishiguro K., Sato K., Omori A., Sato S.,  
RA Shiratsuchi A., Uchida T., Imahori K.;  
RT "A novel brain-specific 25 kDa protein (p25) is phosphorylated by a  
RT Ser/Thr-Pro kinase (TPK II) from tau protein kinase fractions.";  
RL FEBS Lett. 289:37-43(1991).  
FT NON\_TER 1 1  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1375 MW; 2C2822494805D1B7 CRC64;

Query Match 33.3%; Score 3; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SRL 5  
|||  
Db 9 SRL 11

# RESULT 9

Q9QVI3

ID Q9QVI3 PRELIMINARY; PRT; 13 AA.  
AC Q9QVI3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE GP-CS3=CORTICOSTATIC peptide (Fragment).  
OS Cavia (guinea pigs).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae.  
OX NCBI\_TaxID=10140;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92062075; PubMed=1659400;  
RA Hu J., Bennett H.P., Lazure C., Solomon S.;  
RT "Isolation and characterization of corticostatic peptides from guinea  
RT pig bone marrow.";  
RL Biochem. Biophys. Res. Commun. 180:558-565(1991).  
FT NON\_TER 1 1  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1706 MW; D26CD1E63AAE0EB9 CRC64;



Query Match 33.3%; Score 3; DB 11; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLH 6  
|||  
Db 8 RLH 10

RESULT 10

Q85645

ID Q85645 PRELIMINARY; PRT; 13 AA.  
AC Q85645;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Endogenous mouse mammary tumor virus proviral LTR gene product, 5' end  
DE (Fragment).  
OS Mouse mammary tumor virus.  
OC Viruses; Retrovird viruses; Retroviridae; Betaretrovirus.  
OX NCBI\_TaxID=11757;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83139005; PubMed=6298469;  
RA Wheeler D.A., Butel J.S., Medina D., Cardiff R.D., Hager G.L.;  
RT "Transcription of mouse mammary tumor virus: Identification of a  
RT candidate mRNA for the long terminal repeat gene product.";  
RL J. Virol. 46:42-49(1983).  
DR EMBL; M30739; AAA46539.1; -.  
DR InterPro; IPR001213; MMTV\_SAg.  
DR Pfam; PF01054; MMTV\_SAg; 1.  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1686 MW; DA0F2E6B080A7326 CRC64;

Query Match 33.3%; Score 3; DB 15; Length 13;  
Best Local Similarity 100.0%; Pred. No. 4.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSR 4  
|||  
Db 10 NSR 12

RESULT 11

P83159

ID P83159 PRELIMINARY; PRT; 14 AA.  
AC P83159;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Phycobilisome 32.1 kDa linker polypeptide, phycocyanin-associated,  
DE rod (Fragment).  
OS Anabaena sp. (strain L31).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.  
OX NCBI\_TaxID=29412;  
RN [1]  
RP SEQUENCE.

RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;  
 RL Submitted (OCT-2001) to the SWISS-PROT data bank.  
 CC -!- FUNCTION: ROD LINKER PROTEIN, ASSOCIATED WITH PHYCOCYANIN. LINKER  
 CC POLYPEPTIDES DETERMINE THE STATE OF AGGREGATION AND THE LOCATION  
 CC OF THE DISC-SHAPED PHYCOBILIPROTEIN UNITS WITHIN THE PHYCOBILISOME  
 CC AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN ORDER TO MEDIATE A  
 CC DIRECTED AND OPTIMAL ENERGY TRANSFER.  
 CC -!- SUBCELLULAR LOCATION: THIS PROTEIN OCCURS IN THE ROD, IT IS  
 CC ASSOCIATED WITH PHYCOCYANIN (BY SIMILARITY).  
 CC -!- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.  
 KW Phycobilisome; Photosynthesis.  
 FT NON\_TER 14 14  
 SQ SEQUENCE 14 AA; 1405 MW; 96823E44F60A3115 CRC64;

Query Match 33.3%; Score 3; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SRL 5  
 |||  
 Db 7 SRL 9

# RESULT 12

Q16045

ID Q16045 PRELIMINARY; PRT; 14 AA.  
 AC Q16045;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE D3 dopamine receptor (Fragment).  
 GN D3R.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93326145; PubMed=7916609;  
 RA Nagai Y., Ueno S., Saeki Y., Soga F., Yanagihara T.;  
 RT "Expression of the D3 dopamine receptor gene and a novel variant  
 RT transcript generated by alternative splicing in human peripheral blood  
 RT lymphocytes."  
 RL Biochem. Biophys. Res. Commun. 194:368-374 (1993).  
 DR EMBL; S63845; AAB27543.2; -.  
 KW Receptor.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 14 AA; 1586 MW; EA310BEFE94CF1B1 CRC64;

Query Match 33.3%; Score 3; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LHL 7  
 |||  
 Db 10 LHL 12

RESULT 13

Q8TD29

ID Q8TD29 PRELIMINARY; PRT; 14 AA.  
 AC Q8TD29;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE CD8 beta chain (Fragment).  
 GN CD8B1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=11937547;  
 RA Kieffer L.J., Greally J.M., Landres I., Nag S., Nakajima Y.,  
 RA Kohwi-Shigematsu T., Kavathas P.B.;  
 RT "Identification of a Candidate Regulatory Region in the Human CD8 Gene  
 RT Complex by Colocalization of DNase I Hypersensitive Sites and Matrix  
 RT Attachment Regions Which Bind SATB1 and GATA-3.";  
 RL J. Immunol. 168:3915-3922(2002).  
 DR EMBL; AY032722; AAK91556.1; -.  
 FT NON TER 1 1  
 SQ SEQUENCE 14 AA; 1761 MW; 6463EDB22E7ACF3B CRC64;

Query Match 33.3%; Score 3; DB 4; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RLH 6  
 |||  
 Db 2 RLH 4

RESULT 14

Q9QV81

ID Q9QV81 PRELIMINARY; PRT; 14 AA.  
 AC Q9QV81;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Chemoattractant P-1 (Fragment).  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93228656; PubMed=8471066;  
 RA Nakagawa H., Ikesue A., Hatakeyama S., Kato H., Gotoda T.,  
 RA Komorita N., Watanabe K., Miyai H.;  
 RT "Production of an interleukin-8-like chemokine by cytokine-stimulated  
 RT rat NRK-49F fibroblasts and its suppression by anti-inflammatory  
 RT steroids.";

RL Biochem. Pharmacol. 45:1425-1430(1993).

SQ SEQUENCE 14 AA; 1619 MW; 99D3DFA318E6782E CRC64;

Query Match 33.3%; Score 3; DB 11; Length 14;

Best Local Similarity 100.0%; Pred. No. 5.2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LRC 9

|||

Db 3 LRC 5

#### RESULT 15

Q9R545

ID Q9R545 PRELIMINARY; PRT; 15 AA.

AC Q9R545;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE 30S ribosomal protein (Fragment).

OS Mycobacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1765;

RN [1]

RP SEQUENCE.

RX MEDLINE=94009653; PubMed=8405418;

RA Ohara N., Kimura M., Higashi Y., Yamada T.;

RT "Isolation and amino acid sequence of the 30S ribosomal protein S19

RT from Mycobacterium bovis BCG.";

RL FEBS Lett. 331:9-14(1993).

SQ SEQUENCE 15 AA; 1674 MW; 07A36F018AE355A4 CRC64;

Query Match 33.3%; Score 3; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 5.5e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NSR 4

|||

Db 8 NSR 10

Search completed: November 13, 2003, 10:38:06

Job time : 24.125 secs

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:27:36 ; Search time 10.125 Seconds  
(without alignments)  
37.610 Million cell updates/sec

Title: US-09-228-866-1  
Perfect score: 9  
Sequence: 1 CNSRLHLRC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118358

Minimum DB seq length: 7

Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	9	100.0	9	1	US-08-526-710-1
2	9	100.0	9	3	US-08-862-855-1
3	9	100.0	9	3	US-09-226-985-1
4	9	100.0	9	4	US-09-227-906-1
5	5	55.6	9	1	US-08-526-710-5
6	5	55.6	9	3	US-08-862-855-5
7	5	55.6	9	3	US-09-226-985-5
8	5	55.6	9	4	US-09-227-906-5
9	4	44.4	7	1	US-08-137-117D-118
10	4	44.4	7	2	US-08-480-434-78
11	4	44.4	7	2	US-08-436-717-118

12	4	44.4	7	2	US-08-053-451B-78	Sequence 78, Appl
13	4	44.4	7	3	US-08-649-100-13	Sequence 13, Appl
14	4	44.4	7	3	US-08-649-100-29	Sequence 29, Appl
15	4	44.4	8	3	US-09-258-754-256	Sequence 256, App
16	4	44.4	8	3	US-09-042-107-256	Sequence 256, Appl
17	4	44.4	9	3	US-09-258-754-65	Sequence 65, Appl
18	4	44.4	9	3	US-09-042-107-65	Sequence 65, Appl
19	4	44.4	10	2	US-08-488-161-21	Sequence 21, Appl
20	4	44.4	10	3	US-09-273-685-21	Sequence 21, Appl
21	4	44.4	10	5	PCT-US95-11934-21	Sequence 21, Appl
22	4	44.4	11	2	US-08-360-606B-1	Sequence 1, Appli
23	4	44.4	12	3	US-08-747-599A-24	Sequence 24, Appl
24	4	44.4	13	3	US-08-836-075A-156	Sequence 156, App
25	4	44.4	14	5	PCT-US93-06751-99	Sequence 99, Appl
26	4	44.4	15	1	US-08-176-500-31	Sequence 31, Appl
27	4	44.4	15	1	US-08-471-052A-31	Sequence 31, Appl
28	4	44.4	15	1	US-08-189-331-31	Sequence 31, Appl
29	4	44.4	15	2	US-08-471-939-31	Sequence 31, Appl
30	4	44.4	15	2	US-08-471-800-31	Sequence 31, Appl
31	4	44.4	15	2	US-08-488-161-20	Sequence 20, Appl
32	4	44.4	15	2	US-08-471-068-31	Sequence 31, Appl
33	4	44.4	15	3	US-09-273-685-20	Sequence 20, Appl
34	4	44.4	15	5	PCT-US95-11934-20	Sequence 20, Appl
35	4	44.4	16	4	US-09-082-358B-67	Sequence 67, Appl
36	4	44.4	17	4	US-09-082-358B-103	Sequence 103, App
37	4	44.4	17	4	US-09-140-749-14	Sequence 14, Appl
38	4	44.4	18	1	US-07-995-503A-9	Sequence 9, Appli
39	4	44.4	18	1	US-08-390-510-9	Sequence 9, Appli
40	4	44.4	18	1	US-08-390-790-9	Sequence 9, Appli
41	4	44.4	18	2	US-08-390-509-9	Sequence 9, Appli
42	4	44.4	18	3	US-09-149-860A-9	Sequence 9, Appli
43	4	44.4	21	3	US-08-891-640-4	Sequence 4, Appli
44	3	33.3	7	1	US-07-872-644-9	Sequence 9, Appli
45	3	33.3	7	1	US-08-297-494-9	Sequence 9, Appli

#### ALIGNMENTS

##### RESULT 1

US-08-526-710-1

; Sequence 1, Application US/08526710

; Patent No. 5622699

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Pasqualini, Renata

; TITLE OF INVENTION: Method of Identifying Molecules That

; TITLE OF INVENTION: Home to a Selected Organ In Vivo

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-526-710-1

```

```

Query Match          100.0%; Score 9; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CNSRLHLRC 9
        |||||
Db      1 CNSRLHLRC 9

```

## RESULT 2

US-08-862-855-1

```

; Sequence 1, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-1

```

```

Query Match          100.0%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CNSRLHLRC 9
        |||||
Db      1 CNSRLHLRC 9

```

# RESULT 3

US-09-226-985-1

```

; Sequence 1, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```



```

; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-1

```

```

Query Match          100.0%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CNSRLHLRC 9
        |||||
Db      1 CNSRLHLRC 9

```

#### RESULT 4

US-09-227-906-1

```

; Sequence 1, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-1

```

```

Query Match          100.0%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CNSRLHLRC 9
        |||||
Db      1 CNSRLHLRC 9

```

# RESULT 5

US-08-526-710-5

```

; Sequence 5, Application US/08526710
; Patent No. 5622699
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-526-710-5

```

```

Query Match          55.6%; Score 5; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CNSRL 5
        |||||
Db      1 CNSRL 5

```

# RESULT 6

```

US-08-862-855-5
; Sequence 5, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-5

```

```

Query Match          55.6%; Score 5; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CNSRL 5
        |||||
Db      1 CNSRL 5

```

# RESULT 7

US-09-226-985-5

```

; Sequence 5, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-5

```

```

Query Match          55.6%; Score 5; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CNSRL 5
        |||||
Db      1 CNSRL 5

```

# RESULT 8

US-09-227-906-5

```

; Sequence 5, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-5

```

```

Query Match          55.6%; Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CNSRL 5
        |||||
Db      1 CNSRL 5

```

RESULT 9

```

US-08-137-117D-118
; Sequence 118, Application US/08137117D
; Patent No. 5795965
; GENERAL INFORMATION:
; APPLICANT: TSUCHIYA, Masayuki
; APPLICANT: SATO, Koh
; APPLICANT: BENDIG, Mary
; APPLICANT: JONES, Steven
; APPLICANT: SALDANHA, Jose
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR
; NUMBER OF SEQUENCES: 158
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,117D
; FILING DATE: 20-DEC-1993

```

```

; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP92/00544
; FILING DATE: 24-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-32084
; FILING DATE: 19-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-95476
; FILING DATE: 25-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/126/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-137-117D-118

```

```

Query Match          44.4%; Score 4; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 SRLH 6
        ||||
Db      3 SRLH 6

```

# RESULT 10

```

US-08-480-434-78
; Sequence 78, Application US/08480434
; Patent No. 5811248
; GENERAL INFORMATION:
; APPLICANT: Charles C. Ditlow, et al.
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC ANTIGENS,
; TITLE OF INVENTION: ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/480,434
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Albert P. Halluin
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-3660
; TELEFAX: (415) 854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
US-08-480-434-78

```

```

Query Match          44.4%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 SRLH 6
        ||||
Db      3 SRLH 6

```

# RESULT 11

US-08-436-717-118

```

; Sequence 118, Application US/08436717
; Patent No. 5817790
; GENERAL INFORMATION:
; APPLICANT: TSUCHIYA, Masayuki
; APPLICANT: SATO, Koh
; APPLICANT: BENDIG, Mary
; APPLICANT: JONES, Steven
; APPLICANT: SALDANHA, Jose
; TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN
; TITLE OF INVENTION: INTERLEUKIN-6 RECEPTOR
; NUMBER OF SEQUENCES: 158
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```



```

; APPLICATION NUMBER: US/08/436,717
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,117
; FILING DATE: 20-DEC-1993
; APPLICATION NUMBER: WO PCT/JP92/00544
; FILING DATE: 24-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-32084
; FILING DATE: 19-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-95476
; FILING DATE: 25-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/126/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-436-717-118

```

```

Query Match          44.4%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3 SRLH 6
        ||||
Db      3 SRLH 6

```

# RESULT 12

US-08-053-451B-78

```

; Sequence 78, Application US/08053451B
; Patent No. 5955584
; GENERAL INFORMATION:
; APPLICANT: Chen, Francis W.
; APPLICANT: Ditlow, Charles C.
; APPLICANT: Calenoff, Emanuel
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF
; NUMBER OF SEQUENCES: 176
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/053,451B
; FILING DATE: 26-APR-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 7606-033-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; HYPOTHETICAL: N
; ANTI-SENSE: N
US-08-053-451B-78

```

```

Query Match          44.4%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 SRLH 6
        ||||
Db      3 SRLH 6

```

# RESULT 13

US-08-649-100-13

```

; Sequence 13, Application US/08649100
; Patent No. 6114507
; GENERAL INFORMATION:
; APPLICANT: SHIRAKAWA, KAMON
; APPLICANT: MATUSUE, TOMOKAZU
; APPLICANT: NAGATA, SHIGEKAZU
; APPLICANT: CO, MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
; TITLE OF INVENTION: ANTI-FAS LIGAND ANTIBODY AND ASSAY
; TITLE OF INVENTION: METHOD USING THE ANTI-FAS LIGAND ANTIBODY
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/649,100
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1110-160
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-649-100-13

```

```

Query Match          44.4%; Score 4; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 SRLH 6
        ||||
Db      3 SRLH 6

```

#### RESULT 14

```

US-08-649-100-29
; Sequence 29, Application US/08649100
; Patent No. 6114507
; GENERAL INFORMATION:
; APPLICANT: SHIRAKAWA, KAMON
; APPLICANT: MATUSUE, TOMOKAZU
; APPLICANT: NAGATA, SHIGEKAZU
; APPLICANT: CO, MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
; TITLE OF INVENTION: ANTI-FAS LIGAND ANTIBODY AND ASSAY
; TITLE OF INVENTION: METHOD USING THE ANTI-FAS LIGAND ANTIBODY
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/649,100
;   FILING DATE:
;   CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  MURPHY JR, GERALD M
;   REGISTRATION NUMBER:  28,977
;   REFERENCE/DOCKET NUMBER:  1110-160
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (703) 205-8000
;   TELEFAX:  (703) 205-8050
;   INFORMATION FOR SEQ ID NO:  29:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  7 amino acids
;   TYPE:  amino acid
;   STRANDEDNESS:  single
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
US-08-649-100-29

```

```

Query Match          44.4%;  Score 4;  DB 3;  Length 7;
Best Local Similarity 100.0%;  Pred. No. 2.5e+05;
Matches      4;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      3  SRLH 6
        ||||
Db      3  SRLH 6

```

```

RESULT 15
US-09-258-754-256
; Sequence 256, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Dipeptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 256
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-256

```

```

Query Match          44.4%;  Score 4;  DB 3;  Length 8;

```

Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CNSR 4  
    ||||  
Db 1 CNSR 4

Search completed: November 13, 2003, 10:41:53  
Job time : 11.125 secs

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:27:36 ; Search time 10.125 Seconds  
 (without alignments)  
 37.610 Million cell updates/sec

Title: US-09-228-866-2  
 Perfect score: 9  
 Sequence: 1 CENWWGDVC 9

Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118358

Minimum DB seq length: 7  
 Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query Match	Length			
1	9	100.0	9	1	US-08-526-710-2	Sequence 2, Appli
2	9	100.0	9	3	US-08-862-855-2	Sequence 2, Appli
3	9	100.0	9	3	US-09-226-985-2	Sequence 2, Appli
4	9	100.0	9	4	US-09-227-906-2	Sequence 2, Appli
5	4	44.4	7	1	US-08-487-006-129	Sequence 129, App
6	4	44.4	7	2	US-08-488-659A-129	Sequence 129, App
7	4	44.4	7	5	PCT-US91-08328-33	Sequence 33, Appl
8	4	44.4	8	2	US-08-335-832-15	Sequence 15, Appl
9	4	44.4	8	2	US-08-361-864-32	Sequence 32, Appl
10	4	44.4	8	3	US-09-141-127-9	Sequence 9, Appli
11	4	44.4	8	5	PCT-US91-08328-32	Sequence 32, Appl

12	4	44.4	9	1	US-08-078-176-6	Sequence 6, Appli
13	4	44.4	9	1	US-08-078-176-7	Sequence 7, Appli
14	4	44.4	9	1	US-08-078-175-2	Sequence 2, Appli
15	4	44.4	9	1	US-08-250-789A-71	Sequence 71, Appl
16	4	44.4	9	2	US-08-753-781-8	Sequence 8, Appli
17	4	44.4	9	5	PCT-US91-08328-31	Sequence 31, Appl
18	4	44.4	10	1	US-08-250-789A-66	Sequence 66, Appl
19	4	44.4	10	1	US-08-250-789A-76	Sequence 76, Appl
20	4	44.4	10	1	US-08-250-789A-77	Sequence 77, Appl
21	4	44.4	10	1	US-08-250-789A-78	Sequence 78, Appl
22	4	44.4	10	1	US-08-250-789A-79	Sequence 79, Appl
23	4	44.4	10	1	US-08-250-789A-80	Sequence 80, Appl
24	4	44.4	10	1	US-08-250-789A-81	Sequence 81, Appl
25	4	44.4	10	1	US-08-250-789A-82	Sequence 82, Appl
26	4	44.4	10	1	US-08-250-789A-84	Sequence 84, Appl
27	4	44.4	10	1	US-08-250-789A-85	Sequence 85, Appl
28	4	44.4	10	1	US-08-250-789A-86	Sequence 86, Appl
29	4	44.4	10	1	US-08-250-789A-87	Sequence 87, Appl
30	4	44.4	10	1	US-08-250-789A-88	Sequence 88, Appl
31	4	44.4	10	1	US-08-250-789A-89	Sequence 89, Appl
32	4	44.4	10	1	US-08-250-789A-90	Sequence 90, Appl
33	4	44.4	10	1	US-08-433-318A-134	Sequence 134, App
34	4	44.4	10	1	US-08-433-318A-138	Sequence 138, App
35	4	44.4	10	1	US-08-433-318A-142	Sequence 142, App
36	4	44.4	10	1	US-08-433-318A-146	Sequence 146, App
37	4	44.4	10	1	US-08-482-880-27	Sequence 27, Appl
38	4	44.4	10	2	US-08-273-274-27	Sequence 27, Appl
39	4	44.4	10	2	US-08-475-041-27	Sequence 27, Appl
40	4	44.4	10	2	US-08-922-048-134	Sequence 134, App
41	4	44.4	10	2	US-08-922-048-138	Sequence 138, App
42	4	44.4	10	2	US-08-922-048-142	Sequence 142, App
43	4	44.4	10	2	US-08-922-048-146	Sequence 146, App
44	4	44.4	10	2	US-08-335-832-45	Sequence 45, Appl
45	4	44.4	10	2	US-08-484-773-27	Sequence 27, Appl

#### ALIGNMENTS

##### RESULT 1

US-08-526-710-2

; Sequence 2, Application US/08526710

; Patent No. 5622699

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Pasqualini, Renata

; TITLE OF INVENTION: Method of Identifying Molecules That

; TITLE OF INVENTION: Home to a Selected Organ In Vivo

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-526-710-2

```

```

Query Match          100.0%; Score 9; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CENWWGDVC 9
        |||||
Db      1 CENWWGDVC 9

```

# RESULT 2

US-08-862-855-2

```

; Sequence 2, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:

```



```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-2

```

```

Query Match          100.0%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CENWWGDVC 9
        |||||
Db      1 CENWWGDVC 9

```

```

RESULT 3
US-09-226-985-2
; Sequence 2, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-2

```

```

Query Match          100.0%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CENWWGDVC 9
        |||||
Db      1 CENWWGDVC 9

```

#### RESULT 4

```

US-09-227-906-2
; Sequence 2, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-2

```

```

Query Match          100.0%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      1 CENWWGDVC 9
        |||||
Db      1 CENWWGDVC 9

```

# RESULT 5

US-08-487-006-129

```

; Sequence 129, Application US/08487006
; Patent No. 5641861
; GENERAL INFORMATION:
; APPLICANT: Dooley, Colette T.
; APPLICANT: Houghten, Richard A.
; TITLE OF INVENTION: No. 5641861el Mu Opioid Receptor Ligands:
; TITLE OF INVENTION: Agonists and Antagonists
; NUMBER OF SEQUENCES: 222
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/487,006
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-TP 1706
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..7
; OTHER INFORMATION: /note= "With the exception of Gly
; OTHER INFORMATION: in position 6, all amino acids are D-amino acids."
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 7
; OTHER INFORMATION: /note= "Amino acid is amidated at
; OTHER INFORMATION: the C-terminal."
US-08-487-006-129

```

```

Query Match          44.4%; Score 4; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      4 WWGD 7
        ||||
Db      4 WWGD 7

```

# RESULT 6

```

US-08-488-659A-129
; Sequence 129, Application US/08488659A
; Patent No. 5919897
; GENERAL INFORMATION:
; APPLICANT: Dooley, Colette T.
; APPLICANT: Houghten, Richard A.
; TITLE OF INVENTION: Mu Opioid Receptor Ligands:
; TITLE OF INVENTION: Agonists and Antagonists
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,659A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-TP 1705
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..7
; OTHER INFORMATION: /note= "With the exception of Gly
; OTHER INFORMATION: in position 6, all amino acids are D-amino acids."
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 7
; OTHER INFORMATION: /note= "Amino acid is amidated at
; OTHER INFORMATION: the C-terminal."
US-08-488-659A-129

```

```

Query Match          44.4%; Score 4; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 WWGD 7
      ||||
Db      4 WWGD 7

```

```

RESULT 7
PCT-US91-08328-33
; Sequence 33, Application PC/TUS9108328
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Zaverio M.
; APPLICANT: Houghten, Richard A.
; TITLE OF INVENTION: PEPTIDES THAT INHIBIT PLATELET BINDING
; TITLE OF INVENTION: OF ADHESION MOLECULES
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eugene Moroz, MORGAN & FINNEGAN
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  PCT/US91/08328
;   FILING DATE:  19911107
;   CLASSIFICATION:  514
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 07/610,363
;   FILING DATE:  07-NOV-1990
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Moroz, Eugene
;   REGISTRATION NUMBER:  25,237
;   REFERENCE/DOCKET NUMBER:  1198 4079PC
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (212)758-4800
;   TELEFAX:  (212)751-6849
;   TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  33:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  7 amino acids
;   TYPE:  AMINO ACID
;   TOPOLOGY:  unknown
;   MOLECULE TYPE:  peptide
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   FEATURE:
;   NAME/KEY:  Disulfide-bond
;   LOCATION:  1..>7
;   OTHER INFORMATION:  /note= "Sequence linked by
;   OTHER INFORMATION:  interchain disulfide bond at Cys residue with Cys
;   OTHER INFORMATION:  residue on Arg2-Arg-Gly-Asp-Val-Cys"
;   PUBLICATION INFORMATION:
;   DOCUMENT NUMBER:  US 4,683,291
;   FILING DATE:  28-OCT-1985
;   PUBLICATION DATE:  28-JUL-1987
;   PUBLICATION INFORMATION:
;   DOCUMENT NUMBER:  US B1 4,683,291
;   FILING DATE:  28-OCT-1985
;   PUBLICATION DATE:  03-JUL-1990
PCT-US91-08328-33

```

```

Query Match          44.4%;  Score 4;  DB 5;  Length 7;
Best Local Similarity 100.0%;  Pred. No. 2.5e+05;
Matches      4;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      6 GDVC 9
      ||||
Db      4 GDVC 7

```

```

RESULT 8
US-08-335-832-15
; Sequence 15, Application US/08335832
; Patent No. 5925331
; GENERAL INFORMATION:
; APPLICANT:  Dean, Richard T
; APPLICANT:  Lister-James, John

```

```

; TITLE OF INVENTION: Technetium-99m Labeled Peptides for
; TITLE OF INVENTION: Thrombus Imaging
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/335,832
; FILING DATE: 05-JAN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5925331nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 92,216-I
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /label= Acn
; OTHER INFORMATION: /note= "This cysteine residue is blocked at the
; OTHER INFORMATION: sidechain sulfur by covalent linkage to an
; OTHER INFORMATION: acetamido group"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 8
; OTHER INFORMATION: /label= Acn
; OTHER INFORMATION: /note= "This cysteine residue is blocked at the
; OTHER INFORMATION: sidechain sulfur by covalent linkage to an
; OTHER INFORMATION: acetamido group"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 8
; OTHER INFORMATION: /label= Amide
; OTHER INFORMATION: /note= "The carboxyl terminus is modified to an
; OTHER INFORMATION: amide"

```

US-08-335-832-15

Query Match 44.4%; Score 4; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
 ||||  
 Db 3 GDVC 6

RESULT 9  
 US-08-361-864-32  
 ; Sequence 32, Application US/08361864  
 ; Patent No. 5977064  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dean, Richard T  
 ; APPLICANT: Lister-James, John  
 ; TITLE OF INVENTION: Multimeric Polyvalent Antithrombotic  
 ; TITLE OF INVENTION: Agents  
 ; NUMBER OF SEQUENCES: 37  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Allegretti & Witcoff, Ltd.  
 ; STREET: 10 South Wacker Drive, Suite 3000  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/361,864  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/955,466A  
 ; FILING DATE: 19921002  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 5977064nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 92,668  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-715-1000  
 ; TELEFAX: 312-715-1234  
 ; TELEX: 910-221-5317  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 8 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; FEATURE:  
 ; NAME/KEY: Modified-site  
 ; LOCATION: 6..8  
 ; OTHER INFORMATION: /label= Tc-99m-binding  
 ; OTHER INFORMATION: /note= "The sidechain sulfur atom of each cysteine  
 ; OTHER INFORMATION: is protected by an acetamido group, and the  
 ; OTHER INFORMATION: carboxyl-terminal cysteine is an amide"



US-08-361-864-32

Query Match 44.4%; Score 4; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
    ||||  
Db 3 GDVC 6

RESULT 10

US-09-141-127-9

; Sequence 9, Application US/09141127A  
; Patent No. 6083481  
; GENERAL INFORMATION:  
; APPLICANT: Dean, Richard T.  
; APPLICANT: Lister-James, John  
; TITLE OF INVENTION: THROMBUS IMAGING AGENTS  
; FILE REFERENCE: DITI 113.1USC1  
; CURRENT APPLICATION NUMBER: US/09/141,127A  
; CURRENT FILING DATE: 1998-08-27  
; EARLIER APPLICATION NUMBER: 08/335,832  
; EARLIER FILING DATE: 1995-01-05  
; EARLIER APPLICATION NUMBER: PCT/US93/04794  
; EARLIER FILING DATE: 1993-05-21  
; EARLIER APPLICATION NUMBER: 07/886,752  
; EARLIER FILING DATE: 1992-05-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (6)  
; OTHER INFORMATION: BLOCKED:acetamidomethyl  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (8)  
; OTHER INFORMATION: BLOCKED:acetamidomethyl  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (8)  
; OTHER INFORMATION: AMIDATION  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:synthetic  
; OTHER INFORMATION: peptide  
US-09-141-127-9

Query Match 44.4%; Score 4; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
    ||||

## RESULT 11

PCT-US91-08328-32

; Sequence 32, Application PC/TUS9108328  
; GENERAL INFORMATION:  
; APPLICANT: Ruggeri, Zaverio M.  
; APPLICANT: Houghten, Richard A.  
; TITLE OF INVENTION: PEPTIDES THAT INHIBIT PLATELET BINDING  
; TITLE OF INVENTION: OF ADHESION MOLECULES  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eugene Moroz, MORGAN & FINNEGAN  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/08328  
; FILING DATE: 19911107  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/610,363  
; FILING DATE: 07-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Moroz, Eugene  
; REGISTRATION NUMBER: 25,237  
; REFERENCE/DOCKET NUMBER: 1198 4079PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)758-4800  
; TELEFAX: (212)751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: Disulfide-bond  
; LOCATION: 1..>8  
; OTHER INFORMATION: /note= "Sequence linked by  
; OTHER INFORMATION: interchain disulfide bond at Cys residue with Cys  
; OTHER INFORMATION: residue on Arg3-Arg-Gly-Asp-Val-Cys"  
; PUBLICATION INFORMATION:  
; DOCUMENT NUMBER: US 4,683,291  
; FILING DATE: 28-OCT-1985  
; PUBLICATION DATE: 28-JUL-1987

; PUBLICATION INFORMATION:  
; DOCUMENT NUMBER: US B1 4,683,291  
; FILING DATE: 28-OCT-1985  
; PUBLICATION DATE: 03-JUL-1990  
PCT-US91-08328-32

Query Match 44.4%; Score 4; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
|||  
Db 5 GDVC 8

RESULT 12

US-08-078-176-6

; Sequence 6, Application US/08078176  
; Patent No. 5449764  
; GENERAL INFORMATION:  
; APPLICANT: BIRD, Colin R.  
; APPLICANT: RAY, John A.  
; APPLICANT: SCHUCH, Wolfgang W.  
; TITLE OF INVENTION: PLANT DNA  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/078,176  
; FILING DATE: 17-JUN-1993  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9027616.3  
; FILING DATE: 20-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/11372  
; FILING DATE: 19-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kokulis, Paul N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: 200889/SEE36096A/USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-078-176-6

Query Match 44.4%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
|||  
Db 3 CENW 6

RESULT 13

US-08-078-176-7

; Sequence 7, Application US/08078176  
; Patent No. 5449764  
; GENERAL INFORMATION:  
; APPLICANT: BIRD, Colin R.  
; APPLICANT: RAY, John A.  
; APPLICANT: SCHUCH, Wolfgang W.  
; TITLE OF INVENTION: PLANT DNA  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/078,176  
; FILING DATE: 17-JUN-1993  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9027616.3  
; FILING DATE: 20-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/11372  
; FILING DATE: 19-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kokulis, Paul N.  
; REGISTRATION NUMBER: 16,773  
; REFERENCE/DOCKET NUMBER: 200889/SEE36096A/USA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-078-176-7

Query Match 44.4%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CENW 4  
|||  
Db 3 CENW 6

RESULT 14

US-08-078-175-2

; Sequence 2, Application US/08078175

; Patent No. 5484906

; GENERAL INFORMATION:

; APPLICANT: BIRD, Colin R.

; APPLICANT: RAY, John A.

; APPLICANT: SCHUCH, Wolfgang W.

; TITLE OF INVENTION: DNA, DNA CONSTRUCTS, CELLS AND PLANTS

; TITLE OF INVENTION: DERIVED THEREFROM

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cushman, Darby & Cushman

; STREET: 1100 New York Avenue, N.W.

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3918

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/078,175

; FILING DATE: 02-AUG-1993

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9027616.3

; FILING DATE: 20-DEC-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB91/02272

; FILING DATE: 19-DEC-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Kokulis, Paul N.

; REGISTRATION NUMBER: 16,773

; REFERENCE/DOCKET NUMBER: 200621/SEE36096

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-861-3000

; TELEFAX: 202-822-0944

; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-078-175-2

Query Match 44.4%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CENW 4  
|||  
Db 3 CENW 6

RESULT 15

US-08-250-789A-71

; Sequence 71, Application US/08250789A  
; Patent No. 5635597  
; GENERAL INFORMATION:  
; APPLICANT: Barrett, Ronald W.  
; APPLICANT: Chernov-Rogan, Tania  
; APPLICANT: Davis, Ann M.  
; TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors  
; NUMBER OF SEQUENCES: 194  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/250,789A  
; FILING DATE: 27-MAY-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5635597viel, Vernon A.  
; REGISTRATION NUMBER: 32,483  
; REFERENCE/DOCKET NUMBER: 16528A-57/1043  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 71:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-250-789A-71

Query Match 44.4%; Score 4; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGDV 8  
|||  
Db 2 WGDV 5

Search completed: November 13, 2003, 10:41:54  
Job time : 11.125 secs

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:53:11 ; Search time 28.875 Seconds  
(without alignments)  
49.473 Million cell updates/sec

Title: US-09-228-866-2  
Perfect score: 9  
Sequence: 1 CENWWGDVC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 358712

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,



and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9	100.0	9	18	AAW13416	Brain homing pepti
2	9	100.0	9	21	AAB07388	Brain homing pepti
3	9	100.0	9	22	AAE11794	Phage peptide #2 t
4	9	100.0	9	23	AAU10705	Brain homing pepti
5	9	100.0	9	24	ABU59530	Brain receptor tar
6	5	55.6	18	22	ABB45181	Rat albumin-bindin
7	5	55.6	18	22	ABB45191	Rat albumin-bindin
8	4	44.4	7	18	AAW28999	Opioid peptide. S
9	4	44.4	7	18	AAW24360	New peptide which
10	4	44.4	7	19	AAW48546	Integrin receptor
11	4	44.4	7	20	AAAY23111	Opioid peptide whi
12	4	44.4	7	22	AAB75071	Nucleotide-5'-phos
13	4	44.4	8	14	AAR69307	Gp IIB/IIIA recept
14	4	44.4	8	19	AAW48526	Integrin receptor
15	4	44.4	8	19	AAW50590	GPIIb/IIIA recepto
16	4	44.4	8	21	AAAY95459	GPIIb/IIIA recepto
17	4	44.4	8	21	AAAY54973	Peptide ligand for
18	4	44.4	9	19	AAW52082	Targetting peptide
19	4	44.4	10	14	AAR69344	Gp IIB/IIIA recept
20	4	44.4	10	17	AAW16750	p185 binding and e
21	4	44.4	10	17	AAW16738	p185 binding and e
22	4	44.4	10	17	AAW16742	p185 binding and e
23	4	44.4	10	17	AAW16746	p185 binding and e
24	4	44.4	10	18	AAW40684	Peptide which bind
25	4	44.4	10	18	AAW40691	Peptide which bind
26	4	44.4	10	18	AAW40692	Peptide which bind
27	4	44.4	10	18	AAW40693	Peptide which bind
28	4	44.4	10	18	AAW40694	Peptide which bind
29	4	44.4	10	18	AAW40695	Peptide which bind
30	4	44.4	10	18	AAW40696	Peptide which bind
31	4	44.4	10	18	AAW40697	Peptide which bind
32	4	44.4	10	18	AAW40699	Peptide which bind
33	4	44.4	10	18	AAW40700	Peptide which bind
34	4	44.4	10	18	AAW40701	Peptide which bind
35	4	44.4	10	18	AAW40702	Peptide which bind
36	4	44.4	10	18	AAW40703	Peptide which bind
37	4	44.4	10	18	AAW40704	Peptide which bind
38	4	44.4	10	18	AAW40705	Peptide which bind
39	4	44.4	10	21	AAAY54930	Peptide ligand for
40	4	44.4	10	22	AAM42948	Mycoplasma genital
41	4	44.4	11	13	AAR24063	Cell-to-cell bindi
42	4	44.4	11	13	AAR24062	Cell-to-cell bindi
43	4	44.4	11	17	AAW16734	p185 binding and e
44	4	44.4	11	17	AAW16722	p185 binding and e
45	4	44.4	11	17	AAW16726	p185 binding and e

# ALIGNMENTS

RESULT 1

AAW13416

ID AAW13416 standard; Peptide; 9 AA.

XX

AC AAW13416;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;

KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 1997-202359/18.

XX

PT Obtaining compound that homes to selected organ or tissue - by in

PT vivo panning method, specifically to identify brain, kidney,

PT angiogenic vasculature or tumour tissue homing peptide(s)

XX

PS Claim 14; Page 67; 75pp; English.

XX

CC This synthetic peptide is a claimed example of a brain-homing  
 CC peptide that was identified using a novel method for obtaining  
 CC molecules that home to a selected organ or tissue. This in vivo  
 CC panning method typically involves administering a phage display  
 CC library to a subject, and identifying expressed peptides which  
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
 CC vascular tissue or tumour tissue. The isolated peptides (see  
 CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
 CC labels to the selected organ/tissue (claimed) or to identify and/or  
 CC isolate target molecules (claimed). The peptides can be directly  
 CC identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 18; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENWWGDVC 9

Db                    |||||  
1 CENWWGDVC 9

RESULT 2

AAB07388

ID    AAB07388 standard; peptide; 9 AA.

XX

AC    AAB07388;

XX

DT    17-OCT-2000    (first entry)

XX

DE    Brain homing peptide # 2.

XX

KW    Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS    Mus sp.

XX

FH    Key                    Location/Qualifiers

FT    Disulfide-bond    1..9

FT                               /note= "Can optionally form a cyclic peptide"

XX

PN    US6068829-A.

XX

PD    30-MAY-2000.

XX

PF    23-JUN-1997;    97US-0862855.

XX

PR    11-SEP-1995;    95US-0526710.

PR    10-MAR-1997;    97US-0813273.

XX

PA    (BURN-) BURNHAM INST.

XX

PI    Pasqualini R,   Ruoslahti E;

XX

DR    WPI; 2000-410850/35.

XX

PT    Identifying and recovering organ homing molecules or peptides by in  
PT    vivo panning comprises administering a library of diverse peptides  
PT    linked to a tag which facilitates recovery of these peptides    -

XX

PS    Example 2; Column 17; 20pp; English.

XX

CC    The present sequence is a mouse brain homing peptide. This sequence was  
CC    identified by using in vivo panning to screen a library of potential  
CC    organ homing molecules. The present sequence can be used to direct a  
CC    moiety to a the brain tissue, by linking the moiety to the present  
CC    sequence. Examples of potential moieties are drugs, toxins or a  
CC    detectable label.

XX

SQ    Sequence    9 AA;

Query Match                    100.0%;    Score 9;    DB 21;    Length 9;

Best Local Similarity    100.0%;    Pred. No. 9.3e+05;

Matches    9;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy                    1 CENWWGDVC 9

Db

|||||||  
1 CENWWGDVC 9

RESULT 3

AAE11794

ID AAE11794 standard; peptide; 9 AA.

XX

AC AAE11794;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #2 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;

KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo

PT panning that selectively home to a selected organ or tissue useful for

PT treating disease or in diagnostic methods -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENWWGDVC 9  
      |||||  
Db 1 CENWWGDVC 9

RESULT 4

AAU10705

ID AAU10705 standard; peptide; 9 AA.

XX

AC AAU10705;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #2 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
PT identifying molecules that home to a specific organ or tissue, e.g.  
PT identifying a tumour homing molecule to identify the presence of cancer,  
PT by in vivo panning of a library -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
CC home to a selected organ or tissue. The method comprises administering  
CC to the subject the library of diverse molecules, collecting a sample of  
CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
CC the sample several molecules that home to the selected organ or tissue.  
CC The method is useful for identifying molecules, particularly useful for  
CC screening large number of molecules (e.g. peptides), that home to a  
CC specific organ. The identified molecule is useful for e.g. raising an  
CC antibody specific for a target molecule, targeting a desired moiety  
CC (e.g. drug, toxin or detectable label) to the selected organ.  
CC Specifically, the method is useful for identifying the presence of cancer  
CC in a subject by linking an appropriate moiety to a tumour homing  
CC molecule. The present method provides a direct means for identifying

CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 23; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENWWGDVC 9

|||||||

Db 1 CENWWGDVC 9

#### RESULT 5

ABU59530

ID ABU59530 standard; Peptide; 9 AA.

XX

AC ABU59530;

XX

DT 22-APR-2003 (first entry)

XX

DE Brain receptor targeting peptide #2.

XX

KW Targeting ligand; bioactive agent; polymer matrix; cancer; cytostatic;  
KW cathepsin-D substrate; peptides; neuroreceptor; adrenal receptor;  
KW fibronectin; vitronectin; integrin; RGD motif; angiogenic endothelium;  
KW tumour; cationic cancer-targeting peptide.

XX

OS Synthetic.

XX

PN US2002041898-A1.

XX

PD 11-APR-2002.

XX

PF 25-JUL-2001; 2001US-0912609.

XX

PR 05-JAN-2000; 2000US-0478124.

PR 31-OCT-2000; 2000US-0703474.

XX

PA (UNGE/) UNGER E C.

PA (MATS/) MATSUNAGA T O.

PA (RAMA/) RAMASWAMI V.

PA (ROMA/) ROMANOWSKI M J.

XX

PI Unger EC, Matsunaga TO, Ramaswami V, Romanowski MJ;

XX

DR WPI; 2003-208921/20.

XX

PT Targeted delivery system comprising a bioactive agent homogeneously  
PT dispersed in a targeted matrix is especially useful in cancer therapy

PT -

XX

PS Claim 23; Page 37; 46pp; English.

XX

CC The invention relates to a composition comprising a bioactive agent  
CC homogeneously dispersed in a targeted matrix (polymer and targeting  
CC ligand). Also included are a targeted matrix for use as a delivery  
CC vehicle comprising a polymer associated with a targeting ligand,  
CC enhancing the bioavailability of an agent comprising administration  
CC of the composition and treating cancer comprising administration of the  
CC novel composition. The method is useful for targeted delivery of a drug,  
CC especially in cancer therapy. The targeting ligand may be a peptide.  
CC Examples of targeting peptides are disclosed including cathepsin-D  
CC substrate peptides, peptides targeting receptors in the brain and  
CC kidney, peptides recognising fibronectin- and vitronectin-binding  
CC integrins, peptides targeting the RGD (Arg-Gly-Asp)-motif in, e.g.,  
CC antibodies, peptides targeting the angiogenic endothelium of solid  
CC tumours, tissue specific peptides (e.g. of lung, skin, pancreas,  
CC intestine, uterus, adrenal gland and retina), and cationic cancer-  
CC targeting peptides. The present sequence is a peptide targeting  
CC ligand disclosed in the invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 24; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENWWGDVC 9  
|||  
Db 1 CENWWGDVC 9

#### RESULT 6

ABB45181

ID ABB45181 standard; Peptide; 18 AA.

XX

AC ABB45181;

XX

DT 06-FEB-2002 (first entry)

XX

DE Rat albumin-binding peptide #159.

XX

KW Peptide ligand; immunoglobulin G; IgG; serum albumin; rat.

XX

OS Synthetic.

XX

PN WO200145746-A2.

XX

PD 28-JUN-2001.

XX

PF 22-DEC-2000; 2000WO-US35325.

XX

PR 23-DEC-1999; 99US-0173048.

XX

PA (GETH ) GENENTECH INC.

XX

PI Delano WL, Dennis MS, Lowman HB;

XX

DR WPI; 2001-514266/56.

XX

PT Peptide ligands with affinity for immunoglobulin (Ig) G, IgM and/or  
PT human serum albumin which may be conjugated to, and used to prolong the  
PT elimination half-time of active agents from the circulation -

XX

PS Example 6; Fig 8; 71pp; English.

XX

CC The present invention relates to novel peptide ligands with affinity for  
CC immunoglobulin (Ig) G or serum albumin. The peptide ligands may be  
CC conjugated to, and used to prolong the elimination half-time of,  
CC therapeutic and diagnostic agents in the body. The present sequence is a  
CC rat albumin-binding peptide used to illustrate the invention.

XX

SQ Sequence 18 AA;

Query Match 55.6%; Score 5; DB 22; Length 18;

Best Local Similarity 100.0%; Pred. No. 7.6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGDVC 9

|||||

Db 3 WGDVC 7

#### RESULT 7

ABB45191

ID ABB45191 standard; Peptide; 18 AA.

XX

AC ABB45191;

XX

DT 06-FEB-2002 (first entry)

XX

DE Rat albumin-binding peptide #169.

XX

KW Peptide ligand; immunoglobulin G; IgG; serum albumin; rat.

XX

OS Synthetic.

XX

PN WO200145746-A2.

XX

PD 28-JUN-2001.

XX

PF 22-DEC-2000; 2000WO-US35325.

XX

PR 23-DEC-1999; 99US-0173048.

XX

PA (GETH ) GENENTECH INC.

XX

PI Delano WL, Dennis MS, Lowman HB;

XX

DR WPI; 2001-514266/56.

XX

PT Peptide ligands with affinity for immunoglobulin (Ig) G, IgM and/or  
PT human serum albumin which may be conjugated to, and used to prolong the  
PT elimination half-time of active agents from the circulation -

XX



PS Example 6; Fig 8; 71pp; English.

XX

CC The present invention relates to novel peptide ligands with affinity for  
CC immunoglobulin (Ig) G or serum albumin. The peptide ligands may be  
CC conjugated to, and used to prolong the elimination half-time of,  
CC therapeutic and diagnostic agents in the body. The present sequence is a  
CC rat albumin-binding peptide used to illustrate the invention.

XX

SQ Sequence 18 AA;

Query Match 55.6%; Score 5; DB 22; Length 18;

Best Local Similarity 100.0%; Pred. No. 7.6;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGDVC 9

|||||

Db 3 WGDVC 7

#### RESULT 8

AAW28999

ID AAW28999 standard; peptide; 7 AA.

XX

AC AAW28999;

XX

DT 20-JAN-1998 (first entry)

XX

DE Opioid peptide.

XX

KW enkephalin; mu-opioid receptor ligand; agonist; antagonist.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 1

FT /note= "D-form residue"

FT Misc-difference 2

FT /note= "D-form residue"

FT Misc-difference 3

FT /note= "D-form residue"

FT Misc-difference 4

FT /note= "D-form residue"

FT Misc-difference 5

FT /note= "D-form residue"

FT Misc-difference 7

FT /note= "D-form residue"

FT Modified-site 7

FT /note= "the C-terminal is in amide form"

XX

PN US5641861-A.

XX

PD 24-JUN-1997.

XX

PF 07-JUN-1995; 95US-0487006.

XX

PR 07-JUN-1995; 95US-0487006.

XX

PA (TORR-) TORREY PINES INST MOLECULAR STUDIES.  
 XX  
 PI Dooley CT, Houghten RA;  
 XX  
 DR WPI; 1997-340994/31.  
 XX  
 PT New opioid peptides which bind mu receptors specifically - have  
 PT agonist or antagonist activity and are used for study and  
 PT localisation of mu receptors and to treat peripheral side effects of  
 PT morphine etc.  
 XX  
 PS Disclosure; Column 11; 92pp; English.  
 XX  
 CC The patent discloses the following new peptides, which are opioids which  
 CC bind specifically to the mu receptor: Ac-Phe-Arg-Trp-Trp-Tyr-Xaa-NH2 (1);  
 CC Ac-Arg-Trp-Ile-Gly-Trp-Xaa-NH2 (2); Trp-Trp-Pro-Lys-His-Xaa-NH2 (3);  
 CC Trp-Trp-Pro-Xaa1-NH2 (4); Tyr-Pro-Phe-Gly-Phe-Xaa-NH2 (5);  
 CC D-Ile-D-Met-D-Ser-D-Trp-D-Trp-(Gly)n-Xaa2-NH2 (6);  
 CC D-Ile-D-Met-D-Thr-D-Trp-Gly-Xaa2-NH2 (7); Tyr-Al-B2-C3-NH2 (214);  
 CC Pm and red ((Me)x(H)y-Tyr-(NMe)z-Tyr-(Xaa3)z-NH2) (221); and  
 CC Trp-Trp-Pro-D4-(His)z-(Xaa)z-NH2 (222); where Xaa = any natural amino  
 CC acid; Xaa1 = Lys or Arg; n and z = 0 or 1; Xaa2 = Gly or the D form of  
 CC any naturally occurring amino acid; A1 = D-norvaline or D-norleucine;  
 CC B2 = Gly, Phe or Trp; C3 = Trp or naphthylalanine; x and y = 0-2, but  
 CC not over 2 in total; Xaa3 = Phe, DPhe or benzylamino; D4 = Lys or Arg;  
 CC Pm and red indicate permethylation and reduction of all CO in peptide  
 CC links to methylene. These new compounds are useful: (i) for in vitro  
 CC assay and study of opiate receptor subtypes, particularly mu receptors  
 CC in the brain; (ii) for in vivo localisation of receptor subtypes; and  
 CC (iii) therapeutically to block the peripheral effects (e.g. constipation  
 CC and pruritus) of centrally acting pain killers such as morphine.  
 CC They are very selective for the mu opioid receptor, over binding to the  
 CC delta and kappa receptor subtypes.  
 CC The present sequence is a specific example of a peptide (6).  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 18; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWGD 7  
 ||||  
 Db 4 WWGD 7

RESULT 9  
 AAW24360

ID AAW24360 standard; peptide; 7 AA.  
 XX  
 AC AAW24360;  
 XX  
 DT 24-SEP-1997 (first entry)  
 XX  
 DE New peptide which acts as mu-opioid receptor ligand.  
 XX  
 KW mu-receptor; opioid; opiate; agonist; antagonist; diagnosis;

KW analgesic.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1  
FT /note= "D-form residue"  
FT Misc-difference 2  
FT /note= "D-form residue"  
FT Misc-difference 3  
FT /note= "D-form residue"  
FT Misc-difference 4  
FT /note= "D-form residue"  
FT Misc-difference 5  
FT /note= "D-form residue"  
FT Misc-difference 7  
FT /note= "D-form residue, in C-terminal amide form"  
XX  
PN WO9640208-A1.  
XX  
PD 19-DEC-1996.  
XX  
PF 06-JUN-1996; 96WO-US09321.  
XX  
PR 07-JUN-1995; 95US-0476438.  
XX  
PA (TORR-) TORREY PINES INST MOLECULAR STUDIES.  
XX  
PI Dooley CT, Houghten RA;  
XX  
DR WPI; 1997-051895/05.  
XX  
PT New mu opioid receptor binding ligand peptide(s) - useful for  
PT in-vitro and in-vivo diagnosis, as analgesics, and for blocking  
PT peripheral effects of centrally acting drugs, e.g. morphine  
XX  
PS Disclosure; Page 25; 57pp; English.  
XX  
CC The patent discloses eight new groups of opioid peptides which bind  
CC to the mu-receptor to act as agonists or antagonists. The peptides  
CC can be used for in-vitro assays to study opiate receptor subtypes  
CC (especially the mu type) in brain or other tissue samples; and for  
CC in-vivo diagnosis to localise opioid subtypes. The peptides are also  
CC useful as drugs to treat pathologies associated with other compounds  
CC which interact with the opioid receptor system. Therefore they can be  
CC used in medicaments for treating pathologies associated with the mu  
CC receptor and as analgesics. They can be used therapeutically to block  
CC the peripheral effects of centrally acting pain killers, e.g. to  
CC prevent side effects such as constipation and pruritis associated  
CC with morphine. The present sequence represents a specific example  
CC of one of the new groups of peptides, of formula  
CC (D) Ile- (D) Met- (D) Ser- (D) Trp- (D) Trp- Gly- Xaa- NH2 where Xaa = Gly or the  
CC D-form of a naturally occurring amino acid.  
XX  
SQ Sequence 7 AA;

Query Match

44.4%; Score 4; DB 18; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWGD 7  
|||  
Db 4 WWGD 7

RESULT 10

AAW48546

ID AAW48546 standard; peptide; 7 AA.

XX

AC AAW48546;

XX

DT 18-AUG-1998 (first entry)

XX

DE Integrin receptor antagonist peptide 85.

XX

KW Integrin receptor antagonist; cell adhesion modulator; leukocyte;  
KW extracellular matrix; fibronectin; ARDS; thrombosis; inflammation.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Disulfide-bond 1..7

XX

PN US5721210-A.

XX

PD 24-FEB-1998.

XX

PF 07-JUN-1995; 95US-0485019.

XX

PR 04-JUN-1993; 93US-0961889.

PR 09-JUL-1990; 90US-0550330.

PR 09-JUL-1991; 91WO-US04862.

PR 07-JUN-1995; 95US-0485019.

XX

PA (TANA ) TANABE SEIYAKU CO.

XX

PI Cardarelli PM, Chiang S, Lobl TJ;

XX

DR WPI; 1998-168442/15.

XX

PT New cyclic peptide(s) and peptidomimetic compounds - are integrin  
PT receptor antagonists useful in modulating cell adhesion.

XX

PS Example 9; Column 43; 32pp; English.

XX

CC The present sequence represents a synthetic peptide which  
CC acts as an antagonist to integrin receptors. The invention provides  
CC various synthetic peptides which act as cell adhesion modulators because  
CC they mimic extra-cellular matrix ligands or other cell adhesion ligands  
CC that bind to receptors such as integrin receptors, including fibronectin,  
CC laminin, LFA-1, MAC-1, p150,95, vitronectin and gpIIb/IIIa receptors.  
CC Some of the peptides contain the amino acid sequence Arg-Gly-Asp (RGD).  
CC Others contain non-RGD sequences, for e.g RCD sequences, and reverse  
CC orientation forms of amino acid residues. The synthetic peptides

CC are useful in modulating cell adhesion, including adhesion related to  
CC fibronectin, as well as leukocyte adhesion to endothelial cells. They  
CC are also claimed to be useful in the study, diagnosis, treatment or  
CC prevention of diseases which relate to cell adhesion, e.g. adult  
CC respiratory distress syndrome (ARDS), thrombosis and inflammatory  
CC conditions.

XX

SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 19; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9

||||

Db 4 GDVC 7

#### RESULT 11

AAAY23111

ID AAY23111 standard; peptide; 7 AA.

XX

AC AAY23111;

XX

DT 23-AUG-1999 (first entry)

XX

DE Opioid peptide which inhibits binding of enkephalin.

XX

KW Opioid peptide; ligand binding; opioid receptor;

KW micro-selective opioid peptide; enkephalin; opioid receptor system;

KW blocking; peripheral effect; centrally acting pain killer; morphine.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Misc-difference 1..7

FT /note= "with the exception of Gly in position 6,  
FT all amino acids are in the D-form"

FT Modified-site 7

FT /note= "amidated"

XX

PN US5919897-A.

XX

PD 06-JUL-1999.

XX

PF 07-JUN-1995; 95US-0488659.

XX

PR 07-JUN-1995; 95US-0488659.

XX

PA (TORR-) TORREY PINES INST MOLECULAR STUDIES.

XX

PI Dooley CT, Houghten RA;

XX

DR WPI; 1999-394647/33.

XX

PT New opioid peptides useful for blocking the peripheral effects of  
PT centrally acting pain killers such as morphine

XX  
 PS Example 1; Column 11; 92pp; English.  
 XX  
 CC The specification describes opioid peptides, in which each of the  
 CC N atoms in the peptide backbone between respective amino acids is  
 CC modified by permethylation, perallylation, perethylation, perbenzylation  
 CC and pernapthylation. The peptides inhibit ligand binding to an opioid  
 CC receptor. Specifically, the peptides inhibit the micro-selective  
 CC opioid peptide enkephalin. The peptides can be used in vivo  
 CC diagnostically to localize opioid receptor subtypes. They can be used  
 CC to treat pathologies associated with other compounds which interact with  
 CC the opioid receptor system. The peptides are especially useful for  
 CC blocking the peripheral effects of centrally acting pain killers such  
 CC as morphine. The present sequence represents a general structure for  
 CC the opioid peptides of the invention (AAY23092-Y23111 are specific  
 CC examples of the present sequence).  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 20; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWGD 7  
 ||||  
 Db 4 WWGD 7

RESULT 12  
 AAB75071  
 ID AAB75071 standard; Peptide; 7 AA.  
 XX  
 AC AAB75071;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 DE Nucleotide-5'-phosphate producing enzyme variant peptide S72W.  
 XX  
 KW Variant nucleoside-5'-phosphate producing enzyme; mutagenesis;  
 KW transphosphorylation; phosphatase; protein co-ordinate data;  
 KW X-ray structural analysis; three-dimensional structure.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200118184-A1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 01-SEP-2000; 2000WO-JP05973.  
 XX  
 PR 03-SEP-1999; 99JP-0249545.  
 XX  
 PA (AJIN ) AJINOMOTO CO INC.  
 XX  
 PI Ishikawa K, Suzuki E, Gondoh K, Shimba N, Mihara Y, Kawasaki H;  
 PI Kurahashi O, Kouda T, Shimaoka M, Kozutsumi R, Asano Y;  
 XX

DR WPI; 2001-380914/40.  
 XX  
 PT Variant enzyme having elevated nucleoside 5-prime phosphate producing  
 PT activity and having a specific three-dimensional structure for  
 PT production of nucleotides as pharmaceutical intermediates -  
 XX  
 PS Example 11; Fig 7; 150pp; Japanese.  
 XX  
 CC The present invention describes a variant nucleoside-5'-phosphate  
 CC producing enzyme which is a modification of a transphosphorylase or  
 CC phosphatase which contains a lysine, two arginine and two histidine  
 CC residues and in which the C-alpha distances between these residues are  
 CC enclosing a space which permits the binding of a nucleoside with these  
 CC residues. Also described are: (1) a gene encoding the variant enzyme;  
 CC (2) expression vectors containing the DNA; (3) host cells transformed  
 CC by the vectors; (4) preparation of the variant enzyme by culture of the  
 CC transformants; (5) crystals of the enzyme and of a complex of the enzyme  
 CC with molybdenic acid; and (6) selection of inhibitors of acid  
 CC phosphatase or transphosphorylase using the structural coordinates  
 CC derived from the enzyme. The variant enzymes with increased efficiency  
 CC for production of nucleoside 5-phosphates can be used as pharmaceutical  
 CC intermediates. AAH19701 to AAH19785 and AAB75064 to AAB75101 represent  
 CC sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 44.4%; Score 4; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGDV 8  
 ||||  
 Db 4 WGDV 7

# RESULT 13

AAR69307

ID AAR69307 standard; peptide; 8 AA.

XX

AC AAR69307;

XX

DT 25-MAR-2003 (updated)

DT 24-JUN-1995 (first entry)

XX

DE Gp IIb/IIIa receptor ligand used in scintigraphic imaging of thrombi.

XX

KW Scintigraphy; thrombus; thrombi; imaging; specific binding;

KW technetium-99m; radiolabelled; Gp IIb/IIIa receptor ligand.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 6

FT /note= "S-acetamidomethyl-Cys"

FT Modified-site 8

FT /note= "S-acetamidomethyl-Cys-amide"

XX

PN WO9323085-A1.  
 XX  
 PD 25-NOV-1993.  
 XX  
 PF 21-MAY-1993; 93WO-US04794.  
 XX  
 PR 21-MAY-1992; 92US-0886752.  
 XX  
 PA (DIAT-) DIATECH INC.  
 XX  
 PI Dean RT, Lister-James J;  
 XX  
 DR WPI; 1993-386229/48.  
 XX  
 PT Reagent for scintigraphic imaging of thrombi with 99m technetium  
 PT - comprises synthetic peptide which binds to thrombus, covalently  
 PT coupled to metal binding gp.; rapidly cleared from blood and  
 PT tissue  
 XX  
 PS Claim 19; Page 42; 61pp; English.  
 XX  
 CC The invention relates to reagents for scintigraphic imaging of a  
 CC thrombus in-vivo, comprising (A) a specific binding compound capable of  
 CC binding to at least one component of a thrombus, covalently linked  
 CC to (B) a technetium-99m-binding moiety. Specific peptides  
 CC constituting the reagents are claimed as new. The present peptide is  
 CC one such peptide, in which the Cys(Acm)-Gly-Cys(Acm) moiety is the  
 CC 99m-Tc binding moiety and the residue constitutes the thrombus-  
 CC binding component.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 8 AA;  
  
 Query Match 44.4%; Score 4; DB 14; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 6 GDVC 9  
 ||||  
 Db 3 GDVC 6

#### RESULT 14

AAW48526

ID AAW48526 standard; peptide; 8 AA.

XX

AC AAW48526;

XX

DT 18-AUG-1998 (first entry)

XX

DE Integrin receptor antagonist peptide 65.

XX

KW Integrin receptor antagonist; cell adhesion modulator; leukocyte;  
 KW extracellular matrix; fibronectin; ARDS; thrombosis; inflammation.

XX

OS Synthetic.

XX



FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "attached by 1-Adamantaneacetic acid"  
 FT Disulfide-bond 3..8  
 XX  
 PN US5721210-A.  
 XX  
 PD 24-FEB-1998.  
 XX  
 PF 07-JUN-1995; 95US-0485019.  
 XX  
 PR 04-JUN-1993; 93US-0961889.  
 PR 09-JUL-1990; 90US-0550330.  
 PR 09-JUL-1991; 91WO-US04862.  
 PR 07-JUN-1995; 95US-0485019.  
 XX  
 PA (TANA ) TANABE SEIYAKU CO.  
 XX  
 PI Cardarelli PM, Chiang S, Lobl TJ;  
 XX  
 DR WPI; 1998-168442/15.  
 XX  
 PT New cyclic peptide(s) and peptidomimetic compounds - are integrin  
 PT receptor antagonists useful in modulating cell adhesion.  
 XX  
 PS Example 9; Column 42; 32pp; English.  
 XX  
 CC The present sequence represents a synthetic peptide which  
 CC acts as an antagonist to integrin receptors. The invention provides  
 CC various synthetic peptides which act as cell adhesion modulators because  
 CC they mimic extra-cellular matrix ligands or other cell adhesion ligands  
 CC that bind to receptors such as integrin receptors, including fibronectin,  
 CC laminin, LFA-1, MAC-1, p150,95, vitronectin and gpIIb/IIIa receptors.  
 CC Some of the peptides contain the amino acid sequence Arg-Gly-Asp (RGD).  
 CC Others contain non-RGD sequences, for e.g RCD sequences, and reverse  
 CC orientation forms of amino acid residues. The synthetic peptides  
 CC are useful in modulating cell adhesion, including adhesion related to  
 CC fibronectin, as well as leukocyte adhesion to endothelial cells. They  
 CC are also claimed to be useful in the study, diagnosis, treatment or  
 CC prevention of diseases which relate to cell adhesion, e.g. adult  
 CC respiratory distress syndrome (ARDS), thrombosis and inflammatory  
 CC conditions.  
 XX  
 SQ Sequence 8 AA;  
  
 Query Match 44.4%; Score 4; DB 19; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 6 GDVC 9  
 ||||  
 Db 5 GDVC 8

RESULT 15

AAW50590

ID AAW50590 standard; peptide; 8 AA.

XX  
AC AAW50590;  
XX  
DT 16-JUL-1998 (first entry)  
XX  
DE GPIIb/IIIa receptor ligand with technetium-99M binding moiety.  
XX  
KW Technetium-99M label; thrombus imaging; ;GPIIb/IIIa receptor; ligand;  
KW binding moiety.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 6  
FT /note= "Side chain sulphur of Cys blocked by  
FT acetamidomethyl group"  
FT Modified-site 8  
FT /note= "Side chain sulphur of Cys blocked by  
FT acetamidomethyl group, C-terminal amide"  
XX  
PN US5736122-A.  
XX  
PD 07-APR-1998.  
XX  
PF 07-JUN-1995; 95US-0482880.  
XX  
PR 07-JUN-1995; 95US-0482880.  
PR 08-FEB-1991; 91US-0653012.  
PR 27-NOV-1991; 91US-0807062.  
PR 21-MAY-1992; 92US-0886052.  
PR 22-JUN-1994; 94US-0264176.  
PR 11-JUL-1994; 94US-0273274.  
PR 07-JUN-1995; 95US-0480551.  
XX  
PA (DIAT-) DIATIDE INC.  
XX  
PI Dean RT, Lister-James J;  
XX  
DR WPI; 1998-239148/21.  
XX  
PT Reagent useful for preparing thrombus imaging agent - comprises  
PT peptide covalently linked to technetium binding moiety which is  
PT technetium labelled and binds to thrombus component  
XX  
PS Example 2; Columns 11-12; 18pp; English.  
XX  
CC The invention relates to a reagent for preparing a thrombus imaging  
CC agent. It comprises: (a) a specific binding peptide (BP) having a  
CC sequence of 4-100 amino acids; and (b) a technetium-99m binding moiety  
CC (TBM) covalently linked to the peptide which binds to a component of a  
CC thrombus. TBM forms an electrically neutral complex with technetium-99m.  
CC The small size of the peptide component of the reagent enhances blood  
CC and background tissue clearance and it is easy to prepare. This  
CC sequence represents a GPIIb/IIIa receptor binding peptide covalently  
CC linked to a technetium-99m binding moiety.  
XX  
SQ Sequence 8 AA;

Query Match 44.4%; Score 4; DB 19; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
    ||||  
Db 3 GDVC 6

Search completed: November 13, 2003, 10:32:52  
Job time : 29.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:38:27 ; Search time 17.625 Seconds  
(without alignments)  
93.222 Million cell updates/sec

Title: US-09-228-866-2  
Perfect score: 9  
Sequence: 1 CENWWGDVC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 124183

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4	44.4	10	12	US-09-573-822C-257	Sequence 257, App
2	4	44.4	10	14	US-10-100-952-150	Sequence 150, App
3	4	44.4	10	14	US-10-100-952-154	Sequence 154, App
4	4	44.4	10	14	US-10-100-952-158	Sequence 158, App
5	4	44.4	10	14	US-10-100-952-162	Sequence 162, App
6	4	44.4	10	15	US-10-206-699-299	Sequence 299, App
7	4	44.4	11	14	US-10-100-952-134	Sequence 134, App
8	4	44.4	11	14	US-10-100-952-138	Sequence 138, App
9	4	44.4	11	14	US-10-100-952-142	Sequence 142, App
10	4	44.4	11	14	US-10-100-952-146	Sequence 146, App
11	4	44.4	12	10	US-09-952-677-7	Sequence 7, Appli
12	4	44.4	12	15	US-10-284-668-16	Sequence 16, Appl
13	4	44.4	14	12	US-10-254-905-55	Sequence 55, Appl
14	4	44.4	14	14	US-10-100-952-118	Sequence 118, App
15	4	44.4	14	14	US-10-100-952-122	Sequence 122, App
16	4	44.4	14	14	US-10-100-952-126	Sequence 126, App
17	4	44.4	14	14	US-10-100-952-130	Sequence 130, App
18	4	44.4	14	14	US-10-100-952-181	Sequence 181, App
19	4	44.4	14	14	US-10-100-952-185	Sequence 185, App
20	4	44.4	14	14	US-10-100-952-189	Sequence 189, App
21	4	44.4	14	14	US-10-100-952-193	Sequence 193, App
22	4	44.4	14	15	US-10-206-699-5	Sequence 5, Appli
23	4	44.4	15	14	US-10-100-952-101	Sequence 101, App
24	4	44.4	15	14	US-10-100-952-106	Sequence 106, App
25	4	44.4	15	14	US-10-100-952-110	Sequence 110, App
26	4	44.4	15	14	US-10-100-952-114	Sequence 114, App
27	4	44.4	15	14	US-10-100-952-169	Sequence 169, App
28	4	44.4	15	14	US-10-100-952-173	Sequence 173, App
29	4	44.4	15	14	US-10-100-952-177	Sequence 177, App
30	4	44.4	16	12	US-10-125-869A-47	Sequence 47, Appl
31	4	44.4	18	15	US-10-206-699-256	Sequence 256, App
32	4	44.4	18	15	US-10-206-699-262	Sequence 262, App
33	4	44.4	19	12	US-10-378-557-15	Sequence 15, Appl
34	4	44.4	19	12	US-10-378-557-42	Sequence 42, Appl
35	4	44.4	20	12	US-10-280-066-469	Sequence 469, App
36	4	44.4	20	15	US-10-206-699-291	Sequence 291, App
37	4	44.4	20	15	US-10-225-567A-1661	Sequence 1661, Ap
38	3	33.3	7	9	US-09-765-086-218	Sequence 218, App
39	3	33.3	7	9	US-09-989-789-2923	Sequence 2923, Ap
40	3	33.3	7	10	US-09-842-164-12	Sequence 12, Appl
41	3	33.3	7	10	US-09-945-249-50	Sequence 50, Appl
42	3	33.3	7	10	US-09-945-249-58	Sequence 58, Appl
43	3	33.3	7	10	US-09-945-249-61	Sequence 61, Appl
44	3	33.3	7	10	US-09-945-249-63	Sequence 63, Appl

## ALIGNMENTS

## RESULT 1

US-09-573-822C-257

; Sequence 257, Application US/09573822C  
; Publication No. US20030199011A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome sequences  
; FILE REFERENCE: Microbe patent  
; CURRENT APPLICATION NUMBER: US/09/573,822C  
; CURRENT FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 804  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 257  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: mycoplasma genitalium  
; FEATURE:  
; OTHER INFORMATION: Sequence located in MG024 at 351-360 and may interact with Sequence 258  
; OTHER INFORMATION: in this patent.  
US-09-573-822C-257

Query Match 44.4%; Score 4; DB 12; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
|||  
Db 7 GDVC 10

## RESULT 2

US-10-100-952-150

; Sequence 150, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo  
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors  
; FILE REFERENCE: UPN3458  
; CURRENT APPLICATION NUMBER: US/10/100,952  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/111,681  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: 60/076,788  
; PRIOR FILING DATE: 1998-03-04  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 150  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-10-100-952-150

Query Match 44.4%; Score 4; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
|||  
Db 2 CENW 5

#### RESULT 3

US-10-100-952-154  
; Sequence 154, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo  
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors  
; FILE REFERENCE: UPN3458  
; CURRENT APPLICATION NUMBER: US/10/100,952  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/111,681  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: 60/076,788  
; PRIOR FILING DATE: 1998-03-04  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 154  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-10-100-952-154

Query Match 44.4%; Score 4; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
|||  
Db 2 CENW 5

#### RESULT 4

US-10-100-952-158  
; Sequence 158, Application US/10100952  
; Publication No. US20020165193A1

```
; GENERAL INFORMATION:
; APPLICANT: Greene, Mark I.
; APPLICANT: O'Rourke, Donald M.
; APPLICANT: Murali, Ramachandran
; APPLICANT: Park, Byeong Woo
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors
; FILE REFERENCE: UPN3458
; CURRENT APPLICATION NUMBER: US/10/100,952
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/111,681
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/076,788
; PRIOR FILING DATE: 1998-03-04
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-100-952-158
```

```
Query Match          44.4%; Score 4; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 CENW 4
        ||||
Db      2 CENW 5
```

```
RESULT 5
US-10-100-952-162
; Sequence 162, Application US/10100952
; Publication No. US20020165193A1
; GENERAL INFORMATION:
; APPLICANT: Greene, Mark I.
; APPLICANT: O'Rourke, Donald M.
; APPLICANT: Murali, Ramachandran
; APPLICANT: Park, Byeong Woo
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors
; FILE REFERENCE: UPN3458
; CURRENT APPLICATION NUMBER: US/10/100,952
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/111,681
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/076,788
; PRIOR FILING DATE: 1998-03-04
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
```

US-10-100-952-162

Query Match 44.4%; Score 4; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
    ||||  
Db 2 CENW 5

RESULT 6

US-10-206-699-299

; Sequence 299, Application US/10206699  
; Publication No. US20030100510A1  
; GENERAL INFORMATION:  
; APPLICANT: Sundaramoorthy, M.  
; APPLICANT: Hudson, B.  
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NC1 Domain  
Hexamer  
; FILE REFERENCE: MBHB 01-1017  
; CURRENT APPLICATION NUMBER: US/10/206,699  
; CURRENT FILING DATE: 2002-07-26  
; PRIOR APPLICATION NUMBER: US 60/308,523  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: US 60/351,289  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: US 60/366,854  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 60/385,362  
; PRIOR FILING DATE: 2002-06-03  
; NUMBER OF SEQ ID NOS: 307  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 299  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-206-699-299

Query Match 44.4%; Score 4; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDVC 9  
    ||||  
Db 5 GDVC 8

RESULT 7

US-10-100-952-134

; Sequence 134, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo



```
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors
; FILE REFERENCE: UPN3458
; CURRENT APPLICATION NUMBER: US/10/100,952
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/111,681
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/076,788
; PRIOR FILING DATE: 1998-03-04
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 134
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-100-952-134
```

```
Query Match          44.4%; Score 4; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches      4; Conservative    0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 CENW 4
        ||||
Db      3 CENW 6
```

```
RESULT 8
US-10-100-952-138
; Sequence 138, Application US/10100952
; Publication No. US20020165193A1
; GENERAL INFORMATION:
; APPLICANT: Greene, Mark I.
; APPLICANT: O'Rourke, Donald M.
; APPLICANT: Murali, Ramachandran
; APPLICANT: Park, Byeong Woo
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors
; FILE REFERENCE: UPN3458
; CURRENT APPLICATION NUMBER: US/10/100,952
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 09/111,681
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/076,788
; PRIOR FILING DATE: 1998-03-04
; NUMBER OF SEQ ID NOS: 200
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 138
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-100-952-138
```

```
Query Match          44.4%; Score 4; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches      4; Conservative    0; Mismatches      0; Indels      0; Gaps      0;
```

Qy 1 CENW 4  
|||  
Db 3 CENW 6

RESULT 9

US-10-100-952-142

; Sequence 142, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo  
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors  
; FILE REFERENCE: UPN3458  
; CURRENT APPLICATION NUMBER: US/10/100,952  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/111,681  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: 60/076,788  
; PRIOR FILING DATE: 1998-03-04  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 142  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-10-100-952-142

Query Match 44.4%; Score 4; DB 14; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
|||  
Db 3 CENW 6

RESULT 10

US-10-100-952-146

; Sequence 146, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo  
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors  
; FILE REFERENCE: UPN3458  
; CURRENT APPLICATION NUMBER: US/10/100,952  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/111,681  
; PRIOR FILING DATE: 1998-07-08

; PRIOR APPLICATION NUMBER: 60/076,788  
; PRIOR FILING DATE: 1998-03-04  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 146  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-10-100-952-146

Query Match 44.4%; Score 4; DB 14; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
|||  
Db 3 CENW 6

RESULT 11

US-09-952-677-7

; Sequence 7, Application US/09952677  
; Patent No. US20020138876A1  
; GENERAL INFORMATION:  
; APPLICANT: Block, Martina  
; Lorz, Horst  
; Lutticke, Stephanie  
; Walter, Lennart  
; Frohberg, Claus  
; Kossmann, Jens  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ENZYMES  
; FROM WHEAT WHICH ARE INVOLVED IN STARCH  
; SYNTHESIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James F. Haley, Jr., c/o Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/952,677  
; FILING DATE: 14-Sep-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/196,390  
; FILING DATE: 19-No. US20020138876A1-1998  
; APPLICATION NUMBER: DE 196 21 588.9  
; FILING DATE: 29-MAY-1996  
; APPLICATION NUMBER: DE 196 36 917.7

```

; FILING DATE: 11-SEP-1996
; APPLICATION NUMBER: PCT/EP97/02793
; FILING DATE: 28-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: AGREVO-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-952-677-7

```

```

Query Match          44.4%; Score 4; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      6 GDVC 9
        ||||
Db      9 GDVC 12

```

# RESULT 12

US-10-284-668-16

```

; Sequence 16, Application US/10284668
; Publication No. US20030106100A1

```

## GENERAL INFORMATION:

```

; APPLICANT: Kossmann, Jens
; Springer, Franziska
; Abel, Gernot

```

```

; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA

```

## TRANSGENIC

```

; PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES

```

```

; NUMBER OF SEQUENCES: 17

```

## CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020

```

## COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/284,668
; FILING DATE: 29-Oct-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; APPLICATION NUMBER: PCT/EP95/04415
; FILING DATE: 09-NOV-1995
; APPLICATION NUMBER: DE P 44 41 408.0
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Agrevo-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-284-668-16

```

```

Query Match          44.4%; Score 4; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches    4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      6 GDVC 9
        ||||
Db      9 GDVC 12

```

```

RESULT 13
US-10-254-905-55
; Sequence 55, Application US/10254905
; Publication No. US20030186265A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY7,
EXPRESSED HIGHLY IN
; TITLE OF INVENTION: SPINAL CORD
; FILE REFERENCE: D0044 CIP
; CURRENT APPLICATION NUMBER: US/10/254,905
; CURRENT FILING DATE: 2002-09-25
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide

```

US-10-254-905-55

Query Match 44.4%; Score 4; DB 12; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGDV 8  
||||  
Db 3 WGDV 6

RESULT 14

US-10-100-952-118

; Sequence 118, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo  
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors  
; FILE REFERENCE: UPN3458  
; CURRENT APPLICATION NUMBER: US/10/100,952  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/111,681  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: 60/076,788  
; PRIOR FILING DATE: 1998-03-04  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 118  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-10-100-952-118

Query Match 44.4%; Score 4; DB 14; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
||||  
Db 2 CENW 5

RESULT 15

US-10-100-952-122

; Sequence 122, Application US/10100952  
; Publication No. US20020165193A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Mark I.  
; APPLICANT: O'Rourke, Donald M.  
; APPLICANT: Murali, Ramachandran  
; APPLICANT: Park, Byeong Woo  
; TITLE OF INVENTION: Compositions And Methods Of Treating Tumors

; FILE REFERENCE: UPN3458  
; CURRENT APPLICATION NUMBER: US/10/100,952  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: 09/111,681  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: 60/076,788  
; PRIOR FILING DATE: 1998-03-04  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 122  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Sequence  
US-10-100-952-122

Query Match 44.4%; Score 4; DB 14; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CENW 4  
|||  
Db 2 CENW 5

Search completed: November 13, 2003, 11:12:32  
Job time : 18.625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:26:01 ; Search time 8.8125 Seconds  
(without alignments)  
98.215 Million cell updates/sec

Title: US-09-228-866-2  
Perfect score: 9  
Sequence: 1 CENWWGDVC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3752

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	3	33.3	7	2	S57274	triacylglycerol li
2	3	33.3	8	2	A31570	angiotensin-conver
3	3	33.3	11	2	S60354	retinal oxidase -
4	3	33.3	13	2	PH1316	Ig heavy chain DJ
5	3	33.3	14	2	E33098	214K exoantigen (v
6	3	33.3	14	2	S41601	interferon alpha r
7	3	33.3	15	2	PQ0017	terminal protein -
8	3	33.3	15	2	PC4269	unidentified QR310
9	3	33.3	15	2	S14749	3-dehydroquinase -
10	3	33.3	15	4	I38031	hypothetical MN1/T
11	3	33.3	16	2	A42291	tail fiber protein
12	3	33.3	16	2	PC1299	subtilisin (EC 3.4
13	3	33.3	17	2	E40442	integrase homolog
14	3	33.3	18	2	T03799	leader peptide trp
15	3	33.3	18	2	B61110	68K collagen-bindi
16	3	33.3	19	2	S29766	cytochrome c(EDH)
17	3	33.3	20	2	S03987	agglutinin beta-2
18	3	33.3	20	2	PA0022	protein QA100011 -
19	3	33.3	21	2	S68914	ribosomal protein
20	3	33.3	21	2	S03986	agglutinin beta-1
21	2	22.2	7	2	S20446	elastase - Pseudom
22	2	22.2	7	2	S17976	glucose isomerase
23	2	22.2	7	2	B33882	cadmium-binding he
24	2	22.2	7	2	A12016	formylglycinamide
25	2	22.2	7	2	PT0602	T-cell receptor be
26	2	22.2	7	2	PT0628	T-cell receptor be
27	2	22.2	7	2	PT0642	T-cell receptor be
28	2	22.2	7	2	PT0620	T-cell receptor be
29	2	22.2	7	2	PT0666	T-cell receptor be
30	2	22.2	7	2	PT0655	T-cell receptor be
31	2	22.2	7	2	PT0665	T-cell receptor be
32	2	22.2	7	2	PT0556	T-cell receptor be
33	2	22.2	7	2	PT0542	T-cell receptor be
34	2	22.2	7	2	PT0543	T-cell receptor be
35	2	22.2	7	2	PT0544	T-cell receptor be
36	2	22.2	7	2	PT0704	T-cell receptor be
37	2	22.2	7	2	PT0722	T-cell receptor be
38	2	22.2	7	2	PT0567	T-cell receptor be
39	2	22.2	7	2	PT0689	T-cell receptor be
40	2	22.2	7	2	PT0688	T-cell receptor be
41	2	22.2	7	2	PT0586	T-cell receptor be
42	2	22.2	7	2	PT0569	T-cell receptor be



43	2	22.2	7	2	PT0728	T-cell receptor be
44	2	22.2	7	2	B48394	major fat-globule
45	2	22.2	7	2	PD0029	pev-kinin 1 - pena

# ALIGNMENTS

## RESULT 1

S57274

triacylglycerol lipase (EC 3.1.1.3) - *Psychrobacter immobilis* (fragment)

C;Species: *Psychrobacter immobilis*

C;Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 03-Nov-1995

C;Accession: S57274

R;Arpigny, J.L.; Feller, G.; Gerday, C.

Biochim. Biophys. Acta 1263, 103, 1995

A;Title: Corrigendum to "Cloning, sequence and structural features of a lipase from the antarctic facultative psychrophile *Psychrobacter immobilis* B10"

[Biochim. Biophys. Acta 1171 (1993) 331-333].

A;Reference number: S57274; MUID:95359197; PMID:7632728

A;Accession: S57274

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-7 <ARP>

A;Cross-references: EMBL:X67712

C;Keywords: carboxylic ester hydrolase

Query Match	33.3%;	Score 3;	DB 2;	Length 7;
Best Local Similarity	100.0%;	Pred. No. 2.8e+05;		
Matches	3;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY 5 WGD 7

|||

Db 1 WGD 3

## RESULT 2

A31570

angiotensin-converting enzyme inhibitor - yellowfin tuna

C;Species: *Thunnus albacares* (yellowfin tuna)

C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 18-Aug-2000

C;Accession: A31570

R;Kohama, Y.; Matsumoto, S.; Oka, H.; Teramoto, T.; Okabe, M.; Mimura, T.

Biochem. Biophys. Res. Commun. 155, 332-337, 1988

A;Title: Isolation of angiotensin-converting enzyme inhibitor from tuna muscle.

A;Reference number: A31570; MUID:88326322; PMID:3415688

A;Accession: A31570

A;Molecule type: protein

A;Residues: 1-8 <KOH>

A;Note: the source is designated as *Neothunnus macropterus*

C;Superfamily: unassigned animal peptides

C;Keywords: angiotensin-converting enzyme inhibitor

Query Match	33.3%;	Score 3;	DB 2;	Length 8;
Best Local Similarity	100.0%;	Pred. No. 2.8e+05;		
Matches	3;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

Qy            5 WGD 7  
             |||  
Db            6 WGD 8

RESULT 3

S60354

retinal oxidase - rabbit (fragment)

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997

C;Accession: S60354

R;Huang, D.Y.; Ichikawa, Y.

Biochim. Biophys. Acta 1243, 431-436, 1995

A;Title: Identification of essential lysyl and cysteinyl residues, and the amino acid sequence at the substrate-binding site of retinal oxidase.

A;Reference number: S60354; MUID:95244596; PMID:7727518

A;Accession: S60354

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-11 <HUA>

Query Match                    33.3%; Score 3; DB 2; Length 11;  
Best Local Similarity       100.0%; Pred. No. 9.8e+02;  
Matches       3; Conservative    0; Mismatches    0; Indels       0; Gaps       0;

Qy            6 GDV 8  
             |||  
Db            2 GDV 4

RESULT 4

PH1316

Ig heavy chain DJ region (clone C388-107) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

C;Accession: PH1316

R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.  
J. Exp. Med. 176, 1577-1581, 1992

A;Title: Predominance of fetal type DJH joining in young children with B precursor lymphoblastic leukemia as evidence for an in utero transforming event.

A;Reference number: PH1302; MUID:93094761; PMID:1460419

A;Accession: PH1316

A;Molecule type: DNA

A;Residues: 1-13 <WAS>

C;Keywords: heterotetramer; immunoglobulin

Query Match                    33.3%; Score 3; DB 2; Length 13;  
Best Local Similarity       100.0%; Pred. No. 1.1e+03;  
Matches       3; Conservative    0; Mismatches    0; Indels       0; Gaps       0;

Qy            7 DVC 9  
             |||  
Db            4 DVC 6

RESULT 5

E33098

214K exoantigen (version 2) - malaria parasite (Plasmodium falciparum)  
 (fragments)  
 C;Species: Plasmodium falciparum  
 C;Date: 24-Aug-1990 #sequence\_revision 24-Aug-1990 #text\_change 09-Jun-2000  
 C;Accession: E33098  
 R;Nichols, J.H.; Hager, L.P.  
 submitted to the Protein Sequence Database, May 1990  
 A;Reference number: A33098  
 A;Accession: E33098  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-14 <NIC>

Query Match 33.3%; Score 3; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
 |||  
 Db 4 GDV 6

#### RESULT 6

S41601  
 interferon alpha receptor 1 - human (fragments)  
 C;Species: Homo sapiens (man)  
 C;Date: 25-Dec-1994 #sequence\_revision 01-Dec-1995 #text\_change 30-May-1997  
 C;Accession: S41601  
 R;Abramovich, C.; Ratovitski, E.; Lundgren, E.; Revel, M.  
 FEBS Lett. 338, 295-300, 1994  
 A;Title: Identification of mRNAs encoding two different soluble forms of the human interferon alpha-receptor.  
 A;Reference number: S41601; MUID:94139943; PMID:8307198  
 A;Accession: S41601  
 A;Molecule type: mRNA  
 A;Residues: 1-14 <ABR>  
 C;Keywords: cytokine receptor

Query Match 33.3%; Score 3; DB 2; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CEN 3  
 |||  
 Db 6 CEN 8

#### RESULT 7

PQ0017  
 terminal protein - phage M2 (fragment)  
 C;Species: phage M2  
 C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
 C;Accession: PQ0017  
 R;Matsumoto, K.; Takano, H.; Kim, C.I.; Hirokawa, H.  
 Gene 84, 247-255, 1989

A;Title: Primary structure of bacteriophage M2 DNA polymerase: conserved segments within protein-priming DNA polymerases and DNA polymerase I of *Escherichia coli*.

A;Reference number: JQ0161; MUID:90128268; PMID:2515115

A;Accession: PQ0017

A;Molecule type: DNA

A;Residues: 1-15 <MAT>

A;Cross-references: GB:M33144; NID:g215507; PIDN:AAA32367.1; PID:g215508

C;Genetics:

A;Gene: E

C;Superfamily: phage PZA terminal protein

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GDV 8

|||

Db 6 GDV 8

#### RESULT 8

PC4269

unidentified QR310003 protein - rice (fragment)

C;Species: *Oryza sativa* (rice)

C;Date: 28-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 18-Jul-1997

C;Accession: PC4269

R;Kawakami, T.; Kamo, M.; Chen, M.C.; Tsugita, A.  
submitted to JIPID, April 1997

A;Reference number: PC4267

A;Accession: PC4269

A;Molecule type: protein

A;Residues: 1-15 <KAW>

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GDV 8

|||

Db 7 GDV 9

#### RESULT 9

S14749

3-dehydroquinase - *Neurospora crassa* (fragment)

C;Species: *Neurospora crassa*

C;Date: 21-Nov-1993 #sequence\_revision 01-Sep-1995 #text\_change 01-Sep-1995

C;Accession: S14749

R;Chaudhuri, S.; Duncan, K.; Graham, L.D.; Coggins, J.R.  
Biochem. J. 275, 1-6, 1991

A;Title: Identification of the active-site lysine residues of two biosynthetic 3-dehydroquinases.

A;Reference number: S14749; MUID:91207275; PMID:1826831

A;Accession: S14749

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-15 <CHA>

Query Match 33.3%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
|||  
Db 5 GDV 7

RESULT 10

I38031

hypothetical MN1/TEL mutant fusion protein type II - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 20-Apr-2000

C;Accession: I38031

R;Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A.G.; Riegman, P.; Lekanne Deprez, R.; Zwarthoff, E.; Hagemeijer, A.; Grosveld, G.

Oncogene 10, 1511-1519, 1995

A;Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in fusion of the ETS-like TEL gene on 12p13 to the MN1 gene on 22q11.

A;Reference number: I38031; MUID:95249265; PMID:7731705

A;Accession: I38031

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-15 <BUI>

A;Cross-references: EMBL:X85027; NID:g971465; PIDN:CAA59400.1; PID:g971466

C;Comment: This sequence is the chimeric product of a translocation mutation.

C;Genetics:

A;Gene: MN1/ETV6; MN1/TEL

A;Map position: 22q11/12p13

C;Keywords: fusion protein

Query Match 33.3%; Score 3; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
|||  
Db 8 GDV 10

RESULT 11

A42291

tail fiber protein I - phage P2 (fragment)

C;Species: phage P2

C;Date: 10-Jul-1992 #sequence\_revision 10-Jul-1992 #text\_change 30-Sep-1993

C;Accession: A42291

R;Haggard-Ljungquist, E.; Halling, C.; Calendar, R.

J. Bacteriol. 174, 1462-1477, 1992

A;Title: DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages.

A;Reference number: A42291; MUID:92165720; PMID:1531648

A;Accession: A42291

A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-16 <HAG>  
A;Cross-references: GB:M64677

Query Match 33.3%; Score 3; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
|||  
Db 8 GDV 10

#### RESULT 12

PC1299

subtilisin (EC 3.4.21.62) GX - *Bacillus* sp. (strain 6644) (fragment)

C;Species: *Bacillus* sp.

C;Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 11-Nov-1994

C;Accession: PC1299

R;Durham, D.R.

Biochem. Biophys. Res. Commun. 194, 1365-1370, 1993

A;Title: The elastolytic properties of subtilisin GX from alkalophilic *Bacillus* sp. strain 6644 provides a means of differentiation from other subtilisins.

A;Reference number: PC1299; MUID:93356814; PMID:8352796

A;Accession: PC1299

A;Molecule type: protein

A;Residues: 1-16 <DUR>

C;Keywords: hydrolase; serine proteinase

Query Match 33.3%; Score 3; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGD 7  
|||  
Db 5 WGD 7

#### RESULT 13

E40442

integrase homolog (R1 retrotransposable element Nv18) - pteromalid wasp (*Nasonia vitripennis*) (fragment)

C;Species: *Nasonia vitripennis*

C;Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 17-Nov-2000

C;Accession: E40442

R;Jakubczak, J.L.; Burke, W.D.; Eickbush, T.H.

Proc. Natl. Acad. Sci. U.S.A. 88, 3295-3299, 1991

A;Title: Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects.

A;Reference number: A40442; MUID:91195337; PMID:1849649

A;Accession: E40442

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-17 <JAK>

C;Superfamily: silkworm pol protein

Query Match 33.3%; Score 3; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GDV 8  
|||  
Db 4 GDV 6

RESULT 14

T03799

leader peptide trpL - Streptomyces violaceus

N;Alternate names: trp attenuator

C;Species: Streptomyces violaceus

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000

C;Accession: T03799

R;Lin, C.; Paradkar, A.S.; Vining, L.C.

Microbiology 144, 1971-1980, 1998

A;Title: Regulation of an anthranilate synthase gene in Streptomyces venezuelae by a trp attenuator.

A;Reference number: Z15094; MUID:98361043; PMID:9695930

A;Accession: T03799

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-18 <LIN>

A;Cross-references: EMBL:AF012627; NID:g2318111; PID:g2318112

A;Experimental source: strain ISP5230

A;Note: the source is designated as Streptomyces venezuelae

C;Genetics:

A;Gene: trpL

Query Match 33.3%; Score 3; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NWW 5  
|||  
Db 8 NWW 10

RESULT 15

B61110

68K collagen-binding protein, light form - chicken (fragments)

C;Species: Gallus gallus (chicken)

C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-Oct-1994

C;Accession: B61110

R;Tidball, J.G.

J. Biol. Chem. 267, 21211-21219, 1992

A;Title: Identification and distribution of a novel, collagen-binding protein in the developing subepicardium and endomysium.

A;Reference number: A61110; MUID:93016046; PMID:1328225

A;Accession: B61110

A;Molecule type: protein

A;Residues: 1-18 <TID>

C;Keywords: collagen binding

Query Match 33.3%; Score 3; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DVC 9  
 |||  
 Db 1 DVC 3

Search completed: November 13, 2003, 10:39:52  
 Job time : 8.8125 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:55:06 ; Search time 4.875 Seconds  
 (without alignments)  
 86.819 Million cell updates/sec

Title: US-09-228-866-2  
 Perfect score: 9  
 Sequence: 1 CENWWGDVC 9

Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1246

Minimum DB seq length: 7  
 Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	% Query Match Length	DB	ID	Description
1	3	33.3	8	1	ACI_THUAL	P18691 thunnus alb
2	3	33.3	9	1	PPH1_LYCES	P83380 lycopersico
3	3	33.3	9	1	UPA6_HUMAN	P30092 homo sapien
4	3	33.3	15	1	CIQA_RAT	P31720 rattus norv
5	3	33.3	15	1	TERM_BPM2	P19897 bacterioph
6	3	33.3	16	1	FIBA_FELCA	P14450 felis silve
7	3	33.3	17	1	TPIS_PINPS	P81666 pinus pinas



8	3	33.3	20	1	LEC2_MACPO	P18676	maclura pom
9	3	33.3	20	1	LEC3_MACPO	P18677	maclura pom
10	3	33.3	21	1	HCY1_MAISQ	P82302	maia squina
11	3	33.3	21	1	LEC1_MACPO	P18675	maclura pom
12	2	22.2	7	1	FAR1_HELTI	P41871	helisoma tr
13	2	22.2	7	1	PPH2_LYCES	P83379	lycopersico
14	2	22.2	8	1	ACT_CARMA	P80709	carcinus ma
15	2	22.2	8	1	FAR1_PENMO	P83316	penaeus mon
16	2	22.2	8	1	GLUR_HUMAN	P02729	homo sapien
17	2	22.2	8	1	HTF1_PERAM	P04548	periplaneta
18	2	22.2	8	1	HTF2_PERAM	P04549	periplaneta
19	2	22.2	8	1	HTF_TENMO	P25419	tenebrio mo
20	2	22.2	8	1	LCK1_LEUMA	P21140	leucophaea
21	2	22.2	8	1	LCK2_LEUMA	P21141	leucophaea
22	2	22.2	8	1	LCK3_LEUMA	P21142	leucophaea
23	2	22.2	8	1	LCK4_LEUMA	P21143	leucophaea
24	2	22.2	8	1	LCK5_LEUMA	P19987	leucophaea
25	2	22.2	8	1	LCK6_LEUMA	P19988	leucophaea
26	2	22.2	8	1	LCK7_LEUMA	P19989	leucophaea
27	2	22.2	8	1	LCK8_LEUMA	P19990	leucophaea
28	2	22.2	8	1	LMT2_LOCFI	P22396	locusta mig
29	2	22.2	8	1	PLP_BRANA	P81707	brassica na
30	2	22.2	8	1	RT34_BOVIN	P82929	bos taurus
31	2	22.2	9	1	COW_CONVE	P83047	conus ventr
32	2	22.2	9	1	DSIP_RABIT	P01158	oryctolagus
33	2	22.2	9	1	FLA2_TREHY	P80159	treponema h
34	2	22.2	9	1	UHA2_HUMAN	P40929	homo sapien
35	2	22.2	10	1	APE_CAPGI	P80474	capnocytoph
36	2	22.2	10	1	BPP2_BOTIN	P30422	bothrops in
37	2	22.2	10	1	BPP2_BOTJA	P01022	bothrops ja
38	2	22.2	10	1	BPP8_BOTIN	P30426	bothrops in
39	2	22.2	10	1	ESL_LACCA	P81758	lactobacill
40	2	22.2	10	1	FAR7_MACRS	P83280	macrobrachi
41	2	22.2	10	1	FARP_LOCFI	P38553	locusta mig
42	2	22.2	10	1	FARP_MANSE	P18523	manduca sex
43	2	22.2	10	1	FARP_MYTED	P42560	mytilus edu
44	2	22.2	10	1	GLEM_HUMAN	P02728	homo sapien
45	2	22.2	10	1	HTF1_ROMMI	P18110	romalea mic

# ALIGNMENTS

## RESULT 1

ACI\_THUAL

ID ACI\_THUAL STANDARD; PRT; 8 AA.

AC P18691;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-NOV-1990 (Rel. 16, Last annotation update)

DE Angiotensin-converting enzyme inhibitor.

OS Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;

OC Scombridae; Thunnus.

OX NCBI\_TaxID=8236;

RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Muscle;  
 RX MEDLINE=88326322; PubMed=3415688;  
 RA Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;  
 RT "Isolation of angiotensin-converting enzyme inhibitor from tuna  
 RT muscle.";  
 RL Biochem. Biophys. Res. Commun. 155:332-337(1988).  
 DR PIR; A31570; A31570.  
 SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WGD 7  
 |||  
 Db 6 WGD 8

## RESULT 2

PPH1\_LYCES  
 ID PPH1\_LYCES STANDARD; PRT; 9 AA.  
 AC P83380;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Purple acid phosphatase isozyme LeSAP1 (EC 3.1.3.2) (Fragment).  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE, CATALYTIC ACTIVITY, SUBUNIT, SUBCELLULAR LOCATION, AND  
 RP GLYCOSYLATION.  
 RC STRAIN=cv. Moneymaker; TISSUE=Seed;  
 RX MEDLINE=22361242; PubMed=12473124;  
 RA Bozzo G.G., Raghothama K.G., Plaxton W.C.;  
 RT "Purification and characterization of two secreted purple acid  
 RT phosphatase isozymes from phosphate-starved tomato (Lycopersicon  
 RT esculentum) cell cultures.";  
 RL Eur. J. Biochem. 269:6278-6286(2002).  
 CC -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an  
 CC alcohol + phosphate.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PTM: Glycosylated.  
 CC -!- MISCELLANEOUS: In L.esculentum there are at least two isozymes of  
 CC purple acid phosphatase.  
 KW Hydrolase; Glycoprotein.  
 FT NON\_TER 1 1  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1005 MW; 3F17C04B5042CAA8 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
 |||  
 Db 2 GDV 4

RESULT 3

UPA6\_HUMAN

ID UPA6\_HUMAN STANDARD; PRT; 9 AA.  
 AC P30092;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of plasma (Spot 14) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing.";  
 RL Electrophoresis 13:707-714(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 5, ITS MW IS: 48 kDa.  
 DR SWISS-2DPAGE; P30092; HUMAN.  
 FT NON\_TER 1 1  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 935 MW; 5282F2CAA8676447 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
 |||  
 Db 5 GDV 7

RESULT 4

ClQA\_RAT

ID ClQA\_RAT STANDARD; PRT; 15 AA.  
 AC P31720;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Complement Clq subcomponent, A chain (Fragment).  
 GN ClQA.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93218657; PubMed=8464426;  
 RA Wing M.G., Seilly D.J., Bridgman D.J., Harrison R.A.;  
 RT "Rapid isolation and biochemical characterization of rat C1 and C1q.";  
 RL Mol. Immunol. 30:433-440(1993).  
 CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD  
 CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE  
 CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT  
 CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1  
 CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE  
 CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.  
 CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R  
 CC AND S IN THE MOLAR RATION OF 1:2:2. THE C1Q SUBCOMPONENT IS  
 CC COMPOSED OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED  
 CC DIMERS OF THE A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-  
 CC LINKED DIMERS OF THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C  
 CC DIMER BANDS, RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.  
 CC -!- SIMILARITY: Contains 1 collagenous domain.  
 DR InterPro; IPR001073; C1q.  
 DR PROSITE; PS01113; C1Q; PARTIAL.  
 KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;  
 KW Repeat.  
 FT DISULFID 4 4 INTERCHAIN (WITH C-4 IN B CHAIN).  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1488 MW; 1B3D8000B7793965 CRC64;  
  
 Query Match 33.3%; Score 3; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 7 DVC 9  
 |||  
 Db 2 DVC 4

RESULT 5  
 TERM\_BPM2  
 ID TERM\_BPM2 STANDARD; PRT; 15 AA.  
 AC P19897;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE DNA terminal protein (Protein GP3) (Fragment).  
 GN 3 OR E.  
 OS Bacteriophage M2.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
 OC phi-29-like viruses.  
 OX NCBI\_TaxID=10751;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90128268; PubMed=2515115;  
 RA Matsumoto K., Takano H., Kim C.I., Hirokawa H.;  
 RT "Primary structure of bacteriophage M2 DNA polymerase: conserved  
 RT segments within protein-priming DNA polymerases and DNA polymerase I  
 RT of Escherichia coli.";  
 RL Gene 84:247-255(1989).

CC -!- FUNCTION: DNA TERMINAL PROTEIN IS LINKED TO THE 5' ENDS OF  
 CC BOTH STRANDS OF THE GENOME THROUGH A PHOSPHODIESTER BOND BETWEEN  
 CC THE BETA-HYDROXYL GROUP OF A SERINE RESIDUE AND THE 5'-PHOSPHATE  
 CC OF THE TERMINAL DEOXYADENYLATE. THIS PROTEIN IS ESSENTIAL FOR DNA  
 CC REPLICATION AND IS INVOLVED IN THE PRIMING OF DNA ELONGATION.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M33144; AAA32367.1; -.  
 DR PIR; PQ0017; PQ0017.  
 KW Early protein; DNA replication; DNA priming;  
 KW Covalent protein-DNA linkage.  
 FT NON\_TER 1 1  
 FT SITE 5 7 CELL ATTACHMENT SITE (POTENTIAL).  
 SQ SEQUENCE 15 AA; 1797 MW; D3CB AFF8759DEA06 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
 |||  
 Db 6 GDV 8

#### RESULT 6

##### FIBA\_FELCA

ID FIBA\_FELCA STANDARD; PRT; 16 AA.  
 AC P14450;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).  
 GN FGA.  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals."  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES

CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 KW Blood coagulation; Plasma.  
 FT PEPTIDE 1 16 FIBRINOPEPTIDE A.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1620 MW; C3C98EB62D6CC7D3 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
 |||  
 Db 1 GDV 3

# RESULT 7

## TPIS\_PINPS

ID TPIS\_PINPS STANDARD; PRT; 17 AA.  
 AC P81666;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragments).  
 OS Pinus pinaster (Maritime pine).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=71647;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Needle;  
 RX MEDLINE=99274088; PubMed=10344291;  
 RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,  
 RA Frigerio J.-M., Plomion C.;  
 RT "Separation and characterization of needle and xylem maritime pine  
 RT proteins.";  
 RL Electrophoresis 20:1098-1108(1999).  
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone  
 CC phosphate.  
 CC -!- PATHWAY: Plays an important role in several metabolic pathways.  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- INDUCTION: By water stress.  
 CC -!- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC  
 CC AND PLASTID.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN  
 CC (SPOT N139) IS: 5.9, ITS MW IS: 24 kDa.  
 CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.  
 DR InterPro; IPR000652; Triophos\_ismrse.  
 DR PROSITE; PS00171; TIM; PARTIAL.  
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;  
 KW Pentose shunt.  
 FT NON\_TER 1 1  
 FT NON\_CONS 9 10  
 FT NON\_TER 17 17  
 SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;

Query Match 33.3%; Score 3; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 DVC 9  
|||  
Db 3 DVC 5

RESULT 8

LEC2\_MACPO

ID LEC2\_MACPO STANDARD; PRT; 20 AA.  
AC P18676;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Agglutinin beta-2 chain (MPA).  
OS Maclura pomifera (Osage orange).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Rosales; Moraceae; Maclura.  
OX NCBI\_TaxID=3496;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Seed;  
RX MEDLINE=89206218; PubMed=2705782;  
RA Young N.M., Johnston R.A.Z., Szabo A.G., Watson D.C.;  
RT "Homology of the D-galactose-specific lectins from Artocarpus  
RT integrifolia and Maclura pomifera and the role of an unusual small  
RT polypeptide subunit.";  
RL Arch. Biochem. Biophys. 270:596-603(1989).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=98165814; PubMed=9497359;  
RA Lee X., Thompson A., Zhang Z., Ton-That H., Biesterfeldt J., Ogata C.,  
RA Xu L., Johnston R.A., Young N.M.;  
RT "Structure of the complex of Maclura pomifera agglutinin and the T-  
RT antigen disaccharide, Galbeta1,3GalNAc.";  
RL J. Biol. Chem. 273:6312-6318(1998).  
CC -!- FUNCTION: D-GALACTOSE-SPECIFIC LECTIN, BINDS THE T-ANTIGEN  
CC STRUCTURE GAL-BETA1-3-GALNAC.  
CC -!- SUBUNIT: FORMED OF FOUR ALPHA CHAINS AND FOUR BETA CHAINS.  
CC -!- SIMILARITY: TO THE OTHER AGGLUTININ BETA CHAINS.  
DR PIR; S03987; S03987.  
DR PDB; 1JOT; 16-FEB-99.  
KW Lectin; 3D-structure.  
FT STRAND 9 14  
SQ SEQUENCE 20 AA; 2141 MW; AA3882AD5D6370E0 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WGD 7  
|||  
Db 14 WGD 16

RESULT 9

LEC3\_MACPO

ID LEC3\_MACPO STANDARD; PRT; 20 AA.  
 AC P18677;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-AUG-1991 (Rel. 19, Last annotation update)  
 DE Agglutinin beta-3 chain (MPA).  
 OS Maclura pomifera (Osage orange).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Rosales; Moraceae; Maclura.  
 OX NCBI\_TaxID=3496;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=89206218; PubMed=2705782;  
 RA Young N.M., Johnston R.A.Z., Szabo A.G., Watson D.C.;  
 RT "Homology of the D-galactose-specific lectins from Artocarpus  
 RT integrifolia and Maclura pomifera and the role of an unusual small  
 RT polypeptide subunit.";  
 RL Arch. Biochem. Biophys. 270:596-603(1989).  
 CC -!- FUNCTION: D-GALACTOSE-SPECIFIC LECTIN, BINDS THE T-ANTIGEN  
 CC STRUCTURE GAL-BETA1-3-GALNAC.  
 CC -!- SUBUNIT: FORMED OF FOUR ALPHA CHAINS AND FOUR BETA CHAINS.  
 CC -!- SIMILARITY: TO THE OTHER AGGLUTININ BETA CHAINS.  
 DR HSSP; P18676; 1JOT.  
 KW Lectin.  
 SQ SEQUENCE 20 AA; 2082 MW; AA38811BBD6370E0 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGD 7  
 |||  
 Db 14 WGD 16

RESULT 10

HCY1\_MAISQ

ID HCY1\_MAISQ STANDARD; PRT; 21 AA.  
 AC P82302;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hemocyanin subunit 1 (Fragment).  
 OS Maia squinado (Spiny spider crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Majoidea; Majidae; Maia.  
 OX NCBI\_TaxID=99391;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=99259579; PubMed=10327595;  
 RA Stoeva S., Dolashka P., Hristova R., Genov N., Voelter W.;  
 RT "Subunit composition and N-terminal analysis of arthropod



RT hemocyanins.";  
 RL Comp. Biochem. Physiol. 122B:69-75(1999).  
 CC -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers  
 CC occurring freely dissolved in the hemolymph of many mollusks and  
 CC arthropods.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- TISSUE SPECIFICITY: Hemolymph.  
 CC -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY. HEMOCYANIN  
 CC SUBFAMILY.  
 DR InterPro; IPR000896; Hemocyanin.  
 DR InterPro; IPR002227; Tyrosinase.  
 DR PROSITE; PS00209; HEMOCYANIN\_1; PARTIAL.  
 DR PROSITE; PS00210; HEMOCYANIN\_2; PARTIAL.  
 DR PROSITE; PS00497; TYROSINASE\_1; PARTIAL.  
 DR PROSITE; PS00498; TYROSINASE\_2; PARTIAL.  
 KW Transport; Oxygen transport; Copper; Hemolymph.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2556 MW; 7BAB7795FAD2C8B7 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 7e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
 |||  
 Db 4 GDV 6

# RESULT 11

## LEC1\_MACPO

ID LEC1\_MACPO STANDARD; PRT; 21 AA.  
 AC P18675;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-AUG-1991 (Rel. 19, Last annotation update)  
 DE Agglutinin beta-1 chain (MPA).  
 OS Maclura pomifera (Osage orange).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Rosales; Moraceae; Maclura.  
 OX NCBI\_TaxID=3496;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Seed;  
 RX MEDLINE=89206218; PubMed=2705782;  
 RA Young N.M., Johnston R.A.Z., Szabo A.G., Watson D.C.;  
 RT "Homology of the D-galactose-specific lectins from Artocarpus  
 RT integrifolia and Maclura pomifera and the role of an unusual small  
 RT polypeptide subunit."  
 RL Arch. Biochem. Biophys. 270:596-603(1989).  
 CC -!- FUNCTION: D-GALACTOSE-SPECIFIC LECTIN, BINDS THE T-ANTIGEN  
 CC STRUCTURE GAL-BETA1-3-GALNAC.  
 CC -!- SUBUNIT: FORMED OF FOUR ALPHA CHAINS AND FOUR BETA CHAINS.  
 CC -!- SIMILARITY: TO THE OTHER AGGLUTININ BETA CHAINS.  
 DR PIR; S03986; S03986.  
 DR HSSP; P18676; 1JOT.  
 KW Lectin.

SQ SEQUENCE 21 AA; 2196 MW; AA38811BC1BFD0E0 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGD 7  
|||  
Db 15 WGD 17

RESULT 12

FAR1\_HELTI

ID FAR1\_HELTI STANDARD; PRT; 7 AA.  
AC P41871;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE FMRFamide-like neuropeptide GDPFLRF-amide.  
OS Helisoma trivolvis (Snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
OC Lymnaeidae; Planorbidae; Helisoma.  
OX NCBI\_TaxID=27815;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=94286417; PubMed=7912428;  
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
RT trivolvis.";  
RL Peptides 15:31-36(1994).  
CC -!- FUNCTION: APPEARS TO BE INVOLVED IN OSMOREGULATION BY AFFECTING  
CC THE KIDNEY, MANTLE AND SKIN.  
CC -!- TISSUE SPECIFICITY: KIDNEY, SKIN, MANTLE AND THE HEMOLYMPH.  
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 7 7 AMIDATION.  
SQ SEQUENCE 7 AA; 851 MW; 69D40729D76AA810 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GD 7  
||  
Db 1 GD 2

RESULT 13

PPH2\_LYCES

ID PPH2\_LYCES STANDARD; PRT; 7 AA.  
AC P83379;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Purple acid phosphatase isozyme LeSAP2 (EC 3.1.3.2) (Fragment).

OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE, CATALYTIC ACTIVITY, SUBUNIT, SUBCELLULAR LOCATION, AND  
 RP GLYCOSYLATION.  
 RC STRAIN=cv. Moneymaker; TISSUE=Seed;  
 RX MEDLINE=22361242; PubMed=12473124;  
 RA Bozzo G.G., Raghothama K.G., Plaxton W.C.;  
 RT "Purification and characterization of two secreted purple acid  
 RT phosphatase isozymes from phosphate-starved tomato (Lycopersicon  
 RT esculentum) cell cultures."  
 RL Eur. J. Biochem. 269:6278-6286(2002).  
 CC -!- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an  
 CC alcohol + phosphate.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PTM: Glycosylated.  
 CC -!- MISCELLANEOUS: In L.esculentum there are at least two isozymes of  
 CC purple acid phosphatase.  
 KW Hydrolase; Glycoprotein.  
 FT NON\_TER 1 1  
 FT NON\_TER 7 7  
 SQ SEQUENCE 7 AA; 810 MW; 672AA862C9C729A0 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GD 7  
 ||  
 Db 5 GD 6

# RESULT 14

## ACT\_CARMA

ID ACT\_CARMA STANDARD; PRT; 8 AA.  
 AC P80709;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Actin (Fragment).  
 OS Carcinus maenas (Common shore crab) (Green crab).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
 OC Eubrachyura; Portunoidea; Portunidae; Carcinus.  
 OX NCBI\_TaxID=6759;  
 RN [1]  
 RP SEQUENCE.  
 RA Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,  
 RA Baghdassarian D.;  
 RT "A transaldolase. An enzyme implicated in crab steroidogenesis."  
 RL Endocrine 5:23-32(1996).  
 CC -!- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED  
 CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED

CC IN ALL EUKARYOTIC CELLS.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:  
 CC 6.8, ITS MW IS: 46 kDa.  
 CC -!- SIMILARITY: Belongs to the actin family.  
 DR InterPro; IPR004001; Actin.  
 DR InterPro; IPR004000; Actin\_like.  
 DR PROSITE; PS00406; ACTINS\_1; PARTIAL.  
 DR PROSITE; PS00432; ACTINS\_2; PARTIAL.  
 DR PROSITE; PS01132; ACTINS\_ACT\_LIKE; PARTIAL.  
 KW Structural protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 976 MW; 1424005AB2CAAEB3 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DV 8  
 ||  
 Db 3 DV 4

# RESULT 15

## FAR1\_PENMO

ID FAR1\_PENMO STANDARD; PRT; 8 AA.  
 AC P83316;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE FMRFamide-like neuropeptide FLP1 (GDRNFLRF-amide).  
 OS Penaeus monodon (Penaeid shrimp).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;  
 OC Penaeidae; Penaeus.  
 OX NCBI\_TaxID=6687;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Eyestalk;  
 RX MEDLINE=21956277; PubMed=11959015;  
 RA Sithigorngul P., Pupuem J., Krungkasem C., Longyant S.,  
 RA Chaivisuthangkura P., Sithigorngul W., Petsom A.;  
 RT "Seven novel FMRFamide-like neuropeptide sequences from the eyestalk  
 RT of the giant tiger prawn Penaeus monodon."  
 RL Comp. Biochem. Physiol. 131B:325-337(2002).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- MASS SPECTROMETRY: MW=1024.8; METHOD=MALDI.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 1024 MW; 72D40729C4540AA8 CRC64;

Query Match 22.2%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	6	GD	7
Db	1	GD	2

Search completed: November 13, 2003, 10:33:57  
Job time : 6.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      November 13, 2003, 09:58:36 ; Search time 22.125 Seconds
              (without alignments)
              104.971 Million cell updates/sec
```

```
Title:      US-09-228-866-2
Perfect score: 9
Sequence:   1 CENWWGDVC 9
```

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

```
Minimum DB seq length: 7
Maximum DB seq length: 21
```

Post-processing: Listing first 45 summaries

```
Database :          SPTREMBL_23:*
1:   sp_archaea:*
2:   sp_bacteria:*
3:   sp_fungi:*
4:   sp_human:*
5:   sp_invertebrate:*
6:   sp_mammal:*
7:   sp_mhc:*
8:   sp_organelle:*
9:   sp_phage:*
10:  sp_plant:*
11:  sp_rodent:*
12:  sp_virus:*
13:  sp_vertebrate:*
14:  sp_unclassified:*
15:  sp_rvirus:*
16:  sp_bacteriap:*
17:  sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	4	44.4	21	11	Q61917	Q61917	mus musculu
2	3	33.3	8	5	Q9TWH6	Q9twh6	perinereis
3	3	33.3	9	8	Q94NB1	Q94nb1	microcebus
4	3	33.3	9	8	Q94NB2	Q94nb2	microcebus
5	3	33.3	9	8	Q94NA9	Q94na9	daubentonia
6	3	33.3	9	8	Q94XE6	Q94xe6	tectocoris
7	3	33.3	9	8	Q94NB0	Q94nb0	microcebus
8	3	33.3	9	12	Q89491	Q89491	murine minu
9	3	33.3	10	6	Q8SPN8	Q8spn8	macaca mula
10	3	33.3	12	12	Q88577	Q88577	theiler's e
11	3	33.3	12	12	Q88578	Q88578	theiler's e
12	3	33.3	12	12	Q66202	Q66202	transmissib
13	3	33.3	12	12	Q88579	Q88579	theiler's e
14	3	33.3	12	12	Q88575	Q88575	theiler's e
15	3	33.3	12	12	Q88580	Q88580	theiler's e
16	3	33.3	12	12	Q88582	Q88582	theiler's e
17	3	33.3	12	12	Q88581	Q88581	theiler's e
18	3	33.3	12	12	Q88576	Q88576	theiler's e
19	3	33.3	13	8	O99783	O99783	caprimulgus
20	3	33.3	14	10	Q94IT6	Q94it6	fragaria nu
21	3	33.3	15	4	Q9UCJ8	Q9ucj8	homo sapien
22	3	33.3	15	8	P92076	P92076	euhadra her
23	3	33.3	15	8	Q8SJ19	Q8sj19	phalacrocor
24	3	33.3	16	2	Q9R557	Q9r557	bacillus sp
25	3	33.3	16	10	Q9AUA8	Q9aua8	barbarea vu
26	3	33.3	16	10	Q9S8D6	Q9s8d6	triticum ae
27	3	33.3	16	10	Q9AUA7	Q9aua7	rorippa amp
28	3	33.3	17	8	Q94NE6	Q94ne6	cheirogaleu
29	3	33.3	17	8	Q9G120	Q9g120	eulemur ful
30	3	33.3	17	8	Q9G121	Q9g121	eulemur ful
31	3	33.3	17	8	Q9G116	Q9g116	eulemur ful
32	3	33.3	17	8	Q94NE2	Q94ne2	microcebus
33	3	33.3	17	8	Q94NE3	Q94ne3	microcebus
34	3	33.3	17	8	Q9G117	Q9g117	eulemur ful
35	3	33.3	17	8	Q9B0W1	Q9b0w1	propithecus
36	3	33.3	17	8	Q9B8T3	Q9b8t3	propithecus
37	3	33.3	17	8	Q9B0W0	Q9b0w0	propithecus
38	3	33.3	17	8	Q9B8T7	Q9b8t7	propithecus
39	3	33.3	17	8	Q955R8	Q955r8	daubentonia
40	3	33.3	17	8	Q9B0V8	Q9b0v8	propithecus
41	3	33.3	17	8	Q9G118	Q9g118	eulemur ful
42	3	33.3	17	8	Q9G122	Q9g122	eulemur ful
43	3	33.3	17	8	Q94NE5	Q94ne5	cheirogaleu
44	3	33.3	17	8	Q9B8U1	Q9b8u1	avahi lanig
45	3	33.3	17	8	Q9B0V9	Q9b0v9	propithecus

# ALIGNMENTS

# RESULT 1

Q61917

ID Q61917 PRELIMINARY; PRT; 21 AA.  
 AC Q61917;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Genomic murine leukemia virus (MuLV) related sequence (LTR-gag)  
 DE (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83299237; PubMed=6310506;  
 RA Ou C.Y., Boone L.R., Yang W.K.;  
 RT "A novel sequence segment and other nucleotide structural features in  
 RT the long terminal repeat of a BALB/c mouse genomic leukemia virus-  
 RT related DNA clone.";  
 RL Nucleic Acids Res. 11:5603-5620(1983).  
 DR EMBL; X01616; CAA25763.1; -.  
 DR InterPro; IPR000840; Gag\_MA.  
 DR Pfam; PF01140; Gag\_MA; 1.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2369 MW; B636D4F3FBBB11B2 CRC64;

Query Match 44.4%; Score 4; DB 11; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 WGDV 8  
 ||||  
 Db 16 WGDV 19

# RESULT 2

Q9TWH6

ID Q9TWH6 PRELIMINARY; PRT; 8 AA.  
 AC Q9TWH6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE BIOACTIVE peptide P4=PUTATIVE ESOPHAGEAL NEUROREGULATOR.  
 OS Perinereis vancaurica.  
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;  
 OC Phyllodocida; Nereididae; Perinereis.  
 OX NCBI\_TaxID=6355;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95323338; PubMed=7599979;  
 RA Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,  
 RA Fujita T., Minakata H., Nomoto K.;  
 RT "Isolation and characterization of four novel bioactive peptides from  
 RT a polychaete annelid, Perinereis vancaurica.";  
 RL Comp. Biochem. Physiol. C, Pharmacol. Toxicol. Endocrinol. 110:297-

RL 304(1995).

SQ SEQUENCE 8 AA; 989 MW; 954772CAA87B0B59 CRC64;

Query Match 33.3%; Score 3; DB 5; Length 8;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8

|||

Db 4 GDV 6

### RESULT 3

Q94NB1

ID Q94NB1 PRELIMINARY; PRT; 9 AA.  
AC Q94NB1;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Cytochrome oxidase subunit III (Fragment).  
GN COIII.  
OS Microcebus ravelobensis.  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;  
OC Microcebus.  
OX NCBI\_TaxID=122231;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JP299, and JP301;  
RX MEDLINE=21184272; PubMed=11286490;  
RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;  
RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)  
RT based on mitochondrial DNA sequences.";  
RL Mol. Phylogenet. Evol. 19:45-56(2001).  
DR EMBL; AF224630; AAK70571.1; -.  
DR EMBL; AF224631; AAK70575.1; -.  
KW Mitochondrion.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 33.3%; Score 3; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWG 6

|||

Db 6 WWG 8

### RESULT 4

Q94NB2

ID Q94NB2 PRELIMINARY; PRT; 9 AA.  
AC Q94NB2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)



DE Cytochrome oxidase subunit III (Fragment).  
 GN COIII.  
 OS Microcebus murinus (Lesser mouse lemur).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;  
 OC Microcebus.  
 OX NCBI\_TaxID=30608;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JP285, JP288, JP289, JP292, JP308, and JP313;  
 RX MEDLINE=21184272; PubMed=11286490;  
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;  
 RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)  
 RT based on mitochondrial DNA sequences.";  
 RL Mol. Phylogenet. Evol. 19:45-56(2001).  
 DR EMBL; AF224624; AAK70547.1; -.  
 DR EMBL; AF224625; AAK70551.1; -.  
 DR EMBL; AF224626; AAK70555.1; -.  
 DR EMBL; AF224627; AAK70559.1; -.  
 DR EMBL; AF224628; AAK70563.1; -.  
 DR EMBL; AF224629; AAK70567.1; -.  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 33.3%; Score 3; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWG 6  
 |||  
 Db 6 WWG 8

# RESULT 5

Q94NA9

ID Q94NA9 PRELIMINARY; PRT; 9 AA.  
 AC Q94NA9;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Cytochrome oxidase subunit III (Fragment).  
 GN COIII.  
 OS Daubentonia madagascariensis (Aye-aye).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Strepsirhini; Daubentoniidae;  
 OC Daubentonia.  
 OX NCBI\_TaxID=31869;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JP119, and JP120;  
 RX MEDLINE=21184272; PubMed=11286490;  
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;  
 RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)  
 RT based on mitochondrial DNA sequences.";

RL Mol. Phylogenet. Evol. 19:45-56(2001).  
DR EMBL; AF224641; AAK70615.1; -.  
DR EMBL; AF224642; AAK70619.1; -.  
KW Mitochondrion.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 33.3%; Score 3; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWG 6  
|||  
Db 6 WWG 8

#### RESULT 6

Q94XE6

ID Q94XE6 PRELIMINARY; PRT; 9 AA.  
AC Q94XE6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Cytochrome c oxidase subunit III (Fragment).  
GN COX3.  
OS Tectocoris diophthalmus (cotton harlequin bug).  
OG Mitochondrion.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;  
OC Panheteroptera; Pentatomomorpha; Pentatomoidea; Pentatomidae;  
OC Tectocoris.  
OX NCBI\_TaxID=159956;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21396409; PubMed=11504862;  
RA Shao R., Campbell N.J.H., Schmidt E.R., Barker S.C.;  
RT "Increased rate of gene rearrangement in the mitochondrial genomes of  
RT three orders of hemipteroid insects."  
RL Mol. Biol. Evol. 18:1828-1832(2001).  
DR EMBL; AF335990; AAK55283.1; -.  
KW Mitochondrion.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 1206 MW; A2C563636B5041A6 CRC64;

Query Match 33.3%; Score 3; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WWG 6  
|||  
Db 6 WWG 8

#### RESULT 7

Q94NB0

ID Q94NB0 PRELIMINARY; PRT; 9 AA.  
AC Q94NB0;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Cytochrome oxidase subunit III (Fragment).  
 GN COIII.  
 OS Microcebus rufus (brown mouse lemur).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;  
 OC Microcebus.  
 OX NCBI\_TaxID=122232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JP309, JP315, JP316, and JP317;  
 RX MEDLINE=21184272; PubMed=11286490;  
 RA Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;  
 RT "Molecular phylogeny of the lemur family Cheirogaleidae (primates)  
 RT based on mitochondrial DNA sequences.";  
 RL Mol. Phylogenet. Evol. 19:45-56(2001).  
 DR EMBL; AF224636; AAK70595.1; -.  
 DR EMBL; AF224637; AAK70599.1; -.  
 DR EMBL; AF224638; AAK70603.1; -.  
 DR EMBL; AF224639; AAK70607.1; -.  
 KW Mitochondrion.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 9 AA; 1160 MW; D5C563636B5045A2 CRC64;

Query Match 33.3%; Score 3; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWG 6  
 |||  
 Db 6 WWG 8

# RESULT 8

Q89491

ID Q89491 PRELIMINARY; PRT; 9 AA.  
 AC Q89491;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE Hypothetical 1.1 kDa protein.  
 OS Murine minute virus (Murine parvovirus).  
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.  
 OX NCBI\_TaxID=10794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LYMPHOTROPIC VARIANT;  
 RX MEDLINE=86115415; PubMed=3502703;  
 RA Astell C.R., Gardiner E.M., Tattersall P.;  
 RT "DNA sequence of the lymphotropic variant of minute virus of mice,  
 RT MVM(i), and comparison with the DNA sequence of the fibrotropic  
 RT prototype strain.";  
 RL J. Virol. 570:656-669(1986).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=MVM(P);  
 RX MEDLINE=83143341; PubMed=6298737;  
 RA Astell C.R., Thomson M., Merchlinsky M., Ward D.C.;  
 RT "The complete DNA sequence of minute virus of mice, an autonomous  
 RT parvovirus.";  
 RL Nucleic Acids Res. 11:999-1018(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MVM(P);  
 RX MEDLINE=86115415; PubMed=3502703;  
 RA Astell C.R., Gardiner E.M., Tattersall P.;  
 RT "DNA sequence of the lymphotropic variant of minute virus of mice,  
 RT MVM(i), and comparison with the DNA sequence of the fibrotropic  
 RT prototype strain.";  
 RL J. Virol. 57:656-669(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MVM(P);  
 RX MEDLINE=87061199; PubMed=3783817;  
 RA Morgan W.R., Ward D.C.;  
 RT "Three splicing patterns are used to excise the small intron common to  
 RT all minute virus of mice RNAs.";  
 RL J. Virol. 60:1170-1174(1986).  
 DR EMBL; M12032; AAA69570.1; -.  
 DR EMBL; J02275; AAA67112.1; -.  
 DR EMBL; V01115; CAA24311.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 9 AA; 1061 MW; C3FD405863637862 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 WWG 6  
 |||  
 Db 4 WWG 6

# RESULT 9

## Q8SPN8

ID Q8SPN8 PRELIMINARY; PRT; 10 AA.  
 AC Q8SPN8;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Solute carrier family 6 member 4 (Fragment).  
 GN SLC6A4.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Norgren R.B. Jr., Zink M.A., Jia Y., Ojeda S.R., Spindel E.R.;  
 RT "Construction of a targeted rhesus macaque microarray.";

RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AY083583; AAM11998.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 10 AA; 1043 MW; 3B07C4473412CAA8 CRC64;

Query Match 33.3%; Score 3; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GDV 8  
|||  
Db 3 GDV 5

#### RESULT 10

Q88577

ID Q88577 PRELIMINARY; PRT; 12 AA.  
AC Q88577;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Theiler's murine encephalomyelitis virus 5' flank and 5' end cds  
DE (Fragment).  
OS Theiler's encephalomyelitis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Cardiovirus.  
OX NCBI\_TaxID=12124;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TO(4);  
RX MEDLINE=92194426; PubMed=1548749;  
RA Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;  
RT "A single base deletion in the 5' noncoding region of Theiler's virus  
RT attenuates neurovirulence."  
RL J. Virol. 66:1951-1958(1992).  
DR EMBL; M80885; AAA73156.1; -.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1321 MW; F79B7D908E7B5871 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DVC 9  
|||  
Db 9 DVC 11

#### RESULT 11

Q88578

ID Q88578 PRELIMINARY; PRT; 12 AA.  
AC Q88578;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Theiler's murine encephalomyelitis virus 5' flank and 5' end cds  
DE (Fragment).

OS Theiler's encephalomyelitis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Cardiovirus.  
 OX NCBI\_TaxID=12124;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TO(B15);  
 RX MEDLINE=92194426; PubMed=1548749;  
 RA Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;  
 RT "A single base deletion in the 5' noncoding region of Theiler's virus  
 RT attenuates neurovirulence."  
 RL J. Virol. 66:1951-1958(1992).  
 DR EMBL; M80886; AAA73157.1; -.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1321 MW; F79B7D908E7B5871 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 DVC 9  
 |||  
 Db 9 DVC 11

# RESULT 12

Q66202

ID Q66202 PRELIMINARY; PRT; 12 AA.  
 AC Q66202; 072765;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Nucleocapsid protein (Fragment).  
 OS Transmissible gastroenteritis virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=111149;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PURDUE-115;  
 RX MEDLINE=87224815; PubMed=3035066;  
 RA Laude H., Rasschaert D., Huet J.C.;  
 RT "Sequence and N-terminal processing of the transmembrane protein E1 of  
 RT the coronavirus transmissible gastroenteritis virus."  
 RL J. Gen. Virol. 68:1687-1693(1987).  
 DR EMBL; X05598; CAA29092.1; -.  
 KW Nucleocapsid.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1348 MW; 35A1C53F9BD416D8 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 WGD 7  
 |||  
 Db 10 WGD 12

# RESULT 13

Q88579

ID Q88579 PRELIMINARY; PRT; 12 AA.  
AC Q88579;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Theiler's murine encephalomyelitis virus 5' flank and 5' end cds  
DE (Fragment).  
OS Theiler's encephalomyelitis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Cardiovirus.  
OX NCBI\_TaxID=12124;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VL;  
RX MEDLINE=92194426; PubMed=1548749;  
RA Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;  
RT "A single base deletion in the 5' noncoding region of Theiler's virus  
RT attenuates neurovirulence.";  
RL J. Virol. 66:1951-1958(1992).  
DR EMBL; M80887; AAA73158.1; -.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1321 MW; F79B7D908E7B5871 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DVC 9  
|||  
Db 9 DVC 11

# RESULT 14

Q88575

ID Q88575 PRELIMINARY; PRT; 12 AA.  
AC Q88575;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Theiler's murine encephalomyelitis virus 5' flank and 5' end cds  
DE (Fragment).  
OS Theiler's encephalomyelitis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Cardiovirus.  
OX NCBI\_TaxID=12124;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FA;  
RX MEDLINE=92194426; PubMed=1548749;  
RA Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;  
RT "A single base deletion in the 5' noncoding region of Theiler's virus  
RT attenuates neurovirulence.";  
RL J. Virol. 66:1951-1958(1992).

DR EMBL; M80883; AAA73154.1; -.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1321 MW; F79B7D908E7B5871 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DVC 9  
|||  
Db 9 DVC 11

RESULT 15

Q88580

ID Q88580 PRELIMINARY; PRT; 12 AA.  
AC Q88580;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Theiler's murine encephalomyelitis virus 5' flank and 5' end cds  
DE (Fragment).  
OS Theiler's encephalomyelitis virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Cardiovirus.  
OX NCBI\_TaxID=12124;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Vilyuisk;  
RX MEDLINE=92194426; PubMed=1548749;  
RA Pritchard A.E.; Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;  
RT "A single base deletion in the 5' noncoding region of Theiler's virus  
RT attenuates neurovirulence."  
RL J. Virol. 66:1951-1958(1992).  
DR EMBL; M80888; AAA73159.1; -.  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1321 MW; F79B7D908E7B5871 CRC64;

Query Match 33.3%; Score 3; DB 12; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.8e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DVC 9  
|||  
Db 9 DVC 11

Search completed: November 13, 2003, 10:38:08  
Job time : 24.125 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:53:11 ; Search time 28.875 Seconds  
(without alignments)  
49.473 Million cell updates/sec

Title: US-09-228-866-3  
Perfect score: 9  
Sequence: 1 CLSSRLDAC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 358712

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	9	100.0	9	18	AAW13412	Brain homing pepti
2	9	100.0	9	21	AAB07389	Brain homing pepti
3	9	100.0	9	22	AAE11795	Phage peptide #3 t
4	9	100.0	9	23	AAU10706	Brain homing pepti
5	9	100.0	9	24	ABU59532	Brain receptor tar
6	8	88.9	20	21	AAB26819	Peptidic membrane
7	7	77.8	7	18	AAW11186	Brain homing pepti
8	7	77.8	7	21	AAB26822	Peptidic membrane
9	7	77.8	7	21	AAB12005	Brain homing pepti
10	7	77.8	7	22	AAE11811	Phage peptide #19
11	7	77.8	7	23	AAU10722	Brain homing pepti
12	6	66.7	16	22	AAB70197	Exemplary blocking
13	6	66.7	17	22	AAB70196	Exemplary blocking
14	5	55.6	7	22	AAB49817	Human endostatin p
15	5	55.6	9	22	AAB49868	Human endostatin p
16	5	55.6	10	22	AAG93829	Human complementar
17	5	55.6	17	16	AAR91226	Human papillomavir
18	5	55.6	17	21	AAB09119	Hepatitis GB virus
19	5	55.6	17	22	AAB70192	Xenopus CPEB pepti
20	5	55.6	17	22	AAB70194	Exemplary blocking
21	5	55.6	18	22	AAB70198	Xenopus CPEB pepti
22	5	55.6	21	23	ABG66508	IgE Fcepsilon RI b
23	4	44.4	7	14	AAR37709	Delta14 Ser17 hCNT
24	4	44.4	7	21	AAAY85160	Trehalose-releasin
25	4	44.4	7	22	AAG98766	Human cell death p
26	4	44.4	7	22	AAB49818	Human endostatin p
27	4	44.4	7	23	ABG77737	Targetting peptide
28	4	44.4	7	23	ABP49256	Zinc finger protei
29	4	44.4	7	23	ABB46339	Desmoglein-2 CAR s
30	4	44.4	8	18	AAW13438	Brain homing pepti
31	4	44.4	8	18	AAW26811	Octapeptide epitop
32	4	44.4	8	18	AAW26818	Hepatitis C virus
33	4	44.4	8	18	AAW26807	Octapeptide epitop
34	4	44.4	8	18	AAW26808	Octapeptide epitop
35	4	44.4	8	18	AAW26809	Octapeptide epitop
36	4	44.4	8	18	AAW26810	Octapeptide epitop
37	4	44.4	8	19	AAW64270	mMCP-7 peptide sub
38	4	44.4	8	19	AAW75839	Mouse mast cell pr
39	4	44.4	8	19	AAW75808	Mouse mast cell pr
40	4	44.4	8	21	AAB07398	Brain homing pepti
41	4	44.4	8	22	AAE11804	Phage peptide #12
42	4	44.4	8	22	AAB80680	Human glandular ka
43	4	44.4	8	23	AAU10715	Brain homing pepti
44	4	44.4	8	23	ABB46344	Desmoglein-2 CAR s
45	4	44.4	9	16	AAR79075	Alphav/beta3 integ

# ALIGNMENTS

RESULT 1

AAW13412

ID AAW13412 standard; Peptide; 9 AA.

XX

AC AAW13412;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;

KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 1997-202359/18.

XX

PT Obtaining compound that homes to selected organ or tissue - by in

PT vivo panning method, specifically to identify brain, kidney,

PT angiogenic vasculature or tumour tissue homing peptide(s)

XX

PS Claim 10; Page 67; 75pp; English.

XX

CC This synthetic peptide is a claimed example of a brain-homing  
 CC peptide that was identified using a novel method for obtaining  
 CC molecules that home to a selected organ or tissue. This in vivo  
 CC panning method typically involves administering a phage display  
 CC library to a subject, and identifying expressed peptides which  
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
 CC vascular tissue or tumour tissue. The isolated peptides (see  
 CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
 CC labels to the selected organ/tissue (claimed) or to identify and/or  
 CC isolate target molecules (claimed). The peptides can be directly  
 CC identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 18; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 CLSSRLDAC 9

Db                    |||||  
                      1 CLSSRLDAC 9

RESULT 2

AAB07389

ID    AAB07389 standard; peptide; 9 AA.

XX

AC    AAB07389;

XX

DT    17-OCT-2000    (first entry)

XX

DE    Brain homing peptide # 3.

XX

KW    Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS    Mus sp.

XX

FH    Key                    Location/Qualifiers

FT    Disulfide-bond    1..9

FT                                /note= "Can optionally form a cyclic peptide"

XX

PN    US6068829-A.

XX

PD    30-MAY-2000.

XX

PF    23-JUN-1997;    97US-0862855.

XX

PR    11-SEP-1995;    95US-0526710.

PR    10-MAR-1997;    97US-0813273.

XX

PA    (BURN-) BURNHAM INST.

XX

PI    Pasqualini R,    Ruoslahti E;

XX

DR    WPI; 2000-410850/35.

XX

PT    Identifying and recovering organ homing molecules or peptides by in

PT    vivo panning comprises administering a library of diverse peptides

PT    linked to a tag which facilitates recovery of these peptides    -

XX

PS    Example 2; Column 17; 20pp; English.

XX

CC    The present sequence is a mouse brain homing peptide. This sequence was

CC    identified by using in vivo panning to screen a library of potential

CC    organ homing molecules. The present sequence can be used to direct a

CC    moiety to a the brain tissue, by linking the moiety to the present

CC    sequence. Examples of potential moieties are drugs, toxins or a

CC    detectable label. The present sequence contains a SRL amino acid motif.

XX

SQ    Sequence    9 AA;

Query Match                    100.0%;    Score 9;    DB 21;    Length 9;

Best Local Similarity    100.0%;    Pred. No. 9.3e+05;

Matches    9;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy                    1 CLSSRLDAC 9

Db

|||||||  
1 CLSSRLDAC 9

RESULT 3

AAE11795

ID AAE11795 standard; peptide; 9 AA.

XX

AC AAE11795;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #3 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;

KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

FH Key Location/Qualifiers

FT Domain 4..6

FT /label= SRL\_motif

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo  
PT panning that selectively home to a selected organ or tissue useful for  
PT treating disease or in diagnostic methods -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 22; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLSSRLDAC 9  
      |||||  
Db 1 CLSSRLDAC 9

RESULT 4

AAU10706

ID AAU10706 standard; peptide; 9 AA.

XX

AC AAU10706;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #3 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
PT identifying molecules that home to a specific organ or tissue, e.g.  
PT identifying a tumour homing molecule to identify the presence of cancer,  
PT by in vivo panning of a library -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
CC home to a selected organ or tissue. The method comprises administering  
CC to the subject the library of diverse molecules, collecting a sample of  
CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
CC the sample several molecules that home to the selected organ or tissue.  
CC The method is useful for identifying molecules, particularly useful for  
CC screening large number of molecules (e.g. peptides), that home to a  
CC specific organ. The identified molecule is useful for e.g. raising an  
CC antibody specific for a target molecule, targeting a desired moiety

CC (e.g. drug, toxin or detectable label) to the selected organ.  
CC Specifically, the method is useful for identifying the presence of cancer  
CC in a subject by linking an appropriate moiety to a tumour homing  
CC molecule. The present method provides a direct means for identifying  
CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 23; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CLSSRLDAC 9

|||||||

Db 1 CLSSRLDAC 9

#### RESULT 5

ABU59532

ID ABU59532 standard; Peptide; 9 AA.

XX

AC ABU59532;

XX

DT 22-APR-2003 (first entry)

XX

DE Brain receptor targeting peptide #4.

XX

KW Targeting ligand; bioactive agent; polymer matrix; cancer; cytostatic;

KW cathepsin-D substrate; peptides; neuroreceptor; adrenal receptor;

KW fibronectin; vitronectin; integrin; RGD motif; angiogenic endothelium;

KW tumour; cationic cancer-targeting peptide.

XX

OS Synthetic.

XX

PN US2002041898-A1.

XX

PD 11-APR-2002.

XX

PF 25-JUL-2001; 2001US-0912609.

XX

PR 05-JAN-2000; 2000US-0478124.

PR 31-OCT-2000; 2000US-0703474.

XX

PA (UNGE/) UNGER E C.

PA (MATS/) MATSUNAGA T O.

PA (RAMA/) RAMASWAMI V.

PA (ROMA/) ROMANOWSKI M J.

XX

PI Unger EC, Matsunaga TO, Ramaswami V, Romanowski MJ;

XX

DR WPI; 2003-208921/20.

XX

PT Targeted delivery system comprising a bioactive agent homogeneously  
PT dispersed in a targeted matrix is especially useful in cancer therapy  
PT -

XX

PS Claim 23; Page 37; 46pp; English.

XX

CC The invention relates to a composition comprising a bioactive agent  
CC homogeneously dispersed in a targeted matrix (polymer and targeting  
CC ligand). Also included are a targeted matrix for use as a delivery  
CC vehicle comprising a polymer associated with a targeting ligand,  
CC enhancing the bioavailability of an agent comprising administration  
CC of the composition and treating cancer comprising administration of the  
CC novel composition. The method is useful for targeted delivery of a drug,  
CC especially in cancer therapy. The targeting ligand may be a peptide.  
CC Examples of targeting peptides are disclosed including cathepsin-D  
CC substrate peptides, peptides targeting receptors in the brain and  
CC kidney, peptides recognising fibronectin- and vitronectin-binding  
CC integrins, peptides targeting the RGD (Arg-Gly-Asp)-motif in, e.g.,  
CC antibodies, peptides targeting the angiogenic endothelium of solid  
CC tumours, tissue specific peptides (e.g. of lung, skin, pancreas,  
CC intestine, uterus, adrenal gland and retina), and cationic cancer-  
CC targeting peptides. The present sequence is a peptide targeting  
CC ligand disclosed in the invention.

XX

SQ Sequence 9 AA;

Query Match 100.0%; Score 9; DB 24; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLSSRLDAC 9

|||||||

Db 1 CLSSRLDAC 9

#### RESULT 6

AAB26819

ID AAB26819 standard; peptide; 20 AA.

XX

AC AAB26819;

XX

DT 23-JAN-2001 (first entry)

XX

DE Peptidic membrane binding element.

XX

KW Organ perfusion; transplantation; storage; antiinflammatory;  
KW immunosuppressive; vasotropic; complement activation inhibitor;  
KW allograft rejection; ischaemia reperfusion injury.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "Optionally N-Myristoyl-Gly"

FT Modified-site 20

FT /note= "Optionally S-2-Thiopyridyl-Cys-NH2"

XX



PN WO200053007-A1.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-GB00834.  
 XX  
 PR 10-MAR-1999; 99GB-0005503.  
 XX  
 PA (ADPR-) ADPROTECH LTD.  
 XX  
 PI Smith RAG, Pratt JR, Sacks SH;  
 XX  
 DR WPI; 2000-601920/57.  
 XX  
 PT Preparation for perfusing organ prior to transplantation or storage  
 PT comprises soluble derivative of a soluble polypeptide which comprises  
 PT two heterologous membrane binding elements with low membrane affinity  
 PT -  
 XX  
 PS Example 2; Page 20; 47pp; English.  
 XX  
 CC The present invention relates to formulations and preparations for  
 CC perfusing an organ prior to transplantation or storage. The preparation  
 CC comprises a soluble derivative or a polypeptide, which has two or more  
 CC heterologous membrane binding elements. The membrane binding elements are  
 CC capable of interacting, independently and with thermodynamic additivity,  
 CC with membrane components of the organ exposed to extracellular perfusion  
 CC fluids, and a flush storage solution. The preparation exhibits  
 CC antiinflammatory, immunosuppressive and vasotropic activity and works as  
 CC a complement activation inhibitor and an inhibitor of cytotoxic T  
 CC lymphocyte activity. The preparation is used for preparing an organ prior  
 CC to transplantation or storage and for prevention, treatment or  
 CC amelioration of a disease or disorder associated with inflammation,  
 CC inappropriate complement activation or inappropriate activation of  
 CC coagulant or thrombotic processes prior to, during or after  
 CC transplantation or storage of an organ. The preparation is useful for  
 CC treating hyperacute and acute allograft rejection of transplanted organs  
 CC such as kidney, heart, liver or lungs, ischaemia-reperfusion injury in  
 CC transplanted organs, xenograft rejection and corneal graft rejection. The  
 CC present sequence represents a peptidic membrane binding element used in  
 CC an example of the preparation of the invention.  
 XX  
 SQ Sequence 20 AA;

Query Match 88.9%; Score 8; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.051;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRLDAC 9  
 |||||  
 Db 13 LSSRLDAC 20

RESULT 7  
 AAW11186  
 ID AAW11186 standard; Peptide; 7 AA.  
 XX

AC AAW11186;  
 XX  
 DT 15-JAN-1998 (first entry)  
 XX  
 DE Brain homing peptide.  
 XX  
 KW Brain homing peptide; in vivo panning; screening; phage display;  
 KW drug delivery.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9710507-A1.  
 XX  
 PD 20-MAR-1997.  
 XX  
 PF 10-SEP-1996; 96WO-US14600.  
 XX  
 PR 11-SEP-1995; 95US-0526710.  
 PR 11-SEP-1995; 95US-0526708.  
 XX  
 PA (LJOL-) LA JOLLA CANCER RES FOUND.  
 XX  
 PI Pasqualini R, Ruoslahti E;  
 XX  
 DR WPI; 1997-202359/18.  
 XX  
 PT Obtaining compound that homes to selected organ or tissue - by in  
 PT vivo panning method, specifically to identify brain, kidney,  
 PT angiogenic vasculature or tumour tissue homing peptide(s)  
 XX  
 PS Disclosure; Page 45; 75pp; English.  
 XX  
 CC This synthetic peptide is an example of a brain-homing peptide  
 CC that was identified using a claimed method for obtaining  
 CC molecules that home to a selected organ or tissue. This in vivo  
 CC panning method typically involves administering a phage display  
 CC library to a subject, and identifying expressed peptides which  
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
 CC vascular tissue or tumour tissue. The isolated peptides (see  
 CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
 CC labels to the selected organ/tissue (claimed) or to identify and/or  
 CC isolate target molecules (claimed). The peptides can be directly  
 CC identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo.  
 XX  
 SQ Sequence 7 AA;

Query Match 77.8%; Score 7; DB 18; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRLDA 8  
 |||||  
 Db 1 LSSRLDA 7

RESULT 8

AAB26822

ID AAB26822 standard; peptide; 7 AA.

XX

AC AAB26822;

XX

DT 23-JAN-2001 (first entry)

XX

DE Peptidic membrane binding element.

XX

KW Organ perfusion; transplantation; storage; antiinflammatory;  
KW immunosuppressive; vasotropic; complement activation inhibitor;  
KW allograft rejection; ischaemia reperfusion injury.

XX

OS Synthetic.

XX

PN WO200053007-A1.

XX

PD 14-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-GB00834.

XX

PR 10-MAR-1999; 99GB-0005503.

XX

PA (ADPR-) ADPROTECH LTD.

XX

PI Smith RAG, Pratt JR, Sacks SH;

XX

DR WPI; 2000-601920/57.

XX

PT Preparation for perfusing organ prior to transplantation or storage  
PT comprises soluble derivative of a soluble polypeptide which comprises  
PT two heterologous membrane binding elements with low membrane affinity  
PT -

XX

PS Example 2; Page 20; 47pp; English.

XX

CC The present invention relates to formulations and preparations for  
CC perfusing an organ prior to transplantation or storage. The preparation  
CC comprises a soluble derivative or a polypeptide, which has two or more  
CC heterologous membrane binding elements. The membrane binding elements are  
CC capable of interacting, independently and with thermodynamic additivity,  
CC with membrane components of the organ exposed to extracellular perfusion  
CC fluids, and a flush storage solution. The preparation exhibits  
CC antiinflammatory, immunosuppressive and vasotropic activity and works as  
CC a complement activation inhibitor and an inhibitor of cytotoxic T  
CC lymphocyte activity. The preparation is used for preparing an organ prior  
CC to transplantation or storage and for prevention, treatment or  
CC amelioration of a disease or disorder associated with inflammation,  
CC inappropriate complement activation or inappropriate activation of  
CC coagulant or thrombotic processes prior to, during or after  
CC transplantation or storage of an organ. The preparation is useful for  
CC treating hyperacute and acute allograft rejection of transplanted organs  
CC such as kidney, heart, liver or lungs, ischaemia-reperfusion injury in  
CC transplanted organs, xenograft rejection and corneal graft rejection. The  
CC present sequence represents a peptidic membrane binding element used in  
CC an example of the preparation of the invention.

XX

SQ Sequence 7 AA;

Query Match 77.8%; Score 7; DB 21; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRLDA 8

|||||||

Db 1 LSSRLDA 7

RESULT 9

AAB12005

ID AAB12005 standard; peptide; 7 AA.

XX

AC AAB12005;

XX

DT 17-OCT-2000 (first entry)

XX

DE Brain homing peptide # 19.

XX

KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS Mus sp.

XX

PN US6068829-A.

XX

PD 30-MAY-2000.

XX

PF 23-JUN-1997; 97US-0862855.

XX

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 2000-410850/35.

XX

PT Identifying and recovering organ homing molecules or peptides by in

PT vivo panning comprises administering a library of diverse peptides

PT linked to a tag which facilitates recovery of these peptides -

XX

PS Disclosure; Column 14; 20pp; English.

XX

CC The present sequence is a mouse brain homing peptide. This sequence was

CC identified by using in vivo panning to screen a library of potential

CC organ homing molecules. The present sequence can be used to direct a

CC moiety to a the brain tissue, by linking the moiety to the present

CC sequence. Examples of potential moieties are drugs, toxins or a

CC detectable label.

XX

SQ Sequence 7 AA;

Query Match 77.8%; Score 7; DB 21; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRLDA 8  
|||||||  
Db 1 LSSRLDA 7

RESULT 10

AAE11811

ID AAE11811 standard; peptide; 7 AA.

XX

AC AAE11811;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #19 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;  
KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo  
PT panning that selectively home to a selected organ or tissue useful for  
PT treating disease or in diagnostic methods -

XX

PS Disclosure; Column 14; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 7 AA;

Query Match 77.8%; Score 7; DB 22; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRLDA 8  
      |||||  
Db 1 LSSRLDA 7

RESULT 11

AAU10722

ID AAU10722 standard; peptide; 7 AA.

XX

AC AAU10722;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #19 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
PT identifying molecules that home to a specific organ or tissue, e.g.  
PT identifying a tumour homing molecule to identify the presence of cancer,  
PT by in vivo panning of a library -

XX

PS Disclosure; Column 14; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
CC home to a selected organ or tissue. The method comprises administering  
CC to the subject the library of diverse molecules, collecting a sample of  
CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
CC the sample several molecules that home to the selected organ or tissue.  
CC The method is useful for identifying molecules, particularly useful for  
CC screening large number of molecules (e.g. peptides), that home to a  
CC specific organ. The identified molecule is useful for e.g. raising an  
CC antibody specific for a target molecule, targeting a desired moiety

CC (e.g. drug, toxin or detectable label) to the selected organ.  
CC Specifically, the method is useful for identifying the presence of cancer  
CC in a subject by linking an appropriate moiety to a tumour homing  
CC molecule. The present method provides a direct means for identifying  
CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 7 AA;

Query Match 77.8%; Score 7; DB 23; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRLDA 8  
|||  
Db 1 LSSRLDA 7

#### RESULT 12

AAB70197

ID AAB70197 standard; Protein; 16 AA.

XX

AC AAB70197;

XX

DT 02-MAY-2001 (first entry)

XX

DE Exemplary blocking peptide #4.

XX

KW CPEB; cytoplasmic polyadenylation element binding protein; kinase;

KW Eg2; cancer.

XX

OS Unidentified.

XX

PN WO200107466-A1.

XX

PD 01-FEB-2001.

XX

PF 21-JUL-2000; 2000WO-US20254.

XX

PR 21-JUL-1999; 99US-0144903.

XX

PA (UYMA-) UNIV MASSACHUSETTS.

XX

PI Richter JD, Mendez R;

XX

DR WPI; 2001-182774/18.

XX

PT Novel kinase blocking polypeptides useful for treating cancer,  
PT inhibiting synaptic function and for controlling oocyte maturation or  
PT embryogenesis -

XX

PS Claim 10; Page 21; 26pp; English.

XX

CC The present invention relates to a substantially pure kinase blocking  
CC protein. The invention may be used for inhibiting the activity of  
CC kinase Eg2 and for inhibiting the phosphorylation of cytoplasmic  
CC polyadenylation element binding protein (CPEB), in a cell. The  
CC inhibition is performed in vivo or in vitro. It is also useful  
CC for treating a cancer cell.

XX

SQ Sequence 16 AA;

Query Match 66.7%; Score 6; DB 22; Length 16;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRLDA 8

|||||

Db 2 SSRLDA 7

### RESULT 13

AAB70196

ID AAB70196 standard; Protein; 17 AA.

XX

AC AAB70196;

XX

DT 02-MAY-2001 (first entry)

XX

DE Exemplary blocking peptide #3.

XX

KW CPEB; cytoplasmic polyadenylation element binding protein; kinase;

KW Eg2; cancer.

XX

OS Unidentified.

XX

PN WO200107466-A1.

XX

PD 01-FEB-2001.

XX

PF 21-JUL-2000; 2000WO-US20254.

XX

PR 21-JUL-1999; 99US-0144903.

XX

PA (UYMA-) UNIV MASSACHUSETTS.

XX

PI Richter JD, Mendez R;

XX

DR WPI; 2001-182774/18.

XX

PT Novel kinase blocking polypeptides useful for treating cancer,

PT inhibiting synaptic function and for controlling oocyte maturation or

PT embryogenesis -

XX

PS Disclosure; Page 2; 26pp; English.

XX

CC The present invention relates to a substantially pure kinase blocking

CC protein. The invention may be used for inhibiting the activity of

CC kinase Eg2 and for inhibiting the phosphorylation of cytoplasmic

CC polyadenylation element binding protein (CPEB), in a cell. The



CC inhibition is performed in vivo or in vitro. It is also useful  
CC for treating a cancer cell.

XX

SQ Sequence 17 AA;

Query Match 66.7%; Score 6; DB 22; Length 17;

Best Local Similarity 100.0%; Pred. No. 6;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRLDA 8

|||||

Db 3 SSRLDA 8

#### RESULT 14

AAB49817

ID AAB49817 standard; Peptide; 7 AA.

XX

AC AAB49817;

XX

DT 02-MAR-2001 (first entry)

XX

DE Human endostatin peptide fragment SEQ ID NO: 30.

XX

KW Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken;

KW cancer; inflammation; angiogenesis-dependent disease.

XX

OS Homo sapiens.

XX

PN WO200067771-A1.

XX

PD 16-NOV-2000.

XX

PF 02-MAY-2000; 2000WO-US12063.

XX

PR 06-MAY-1999; 99US-0132907.

PR 14-JUL-1999; 99US-0353333.

XX

PA (BURN-) BURNHAM INST.

XX

PI Vuori K;

XX

DR WPI; 2001-040937/05.

XX

PT Endostatin peptide comprising at least four endostatin amino acid

PT residues are e.g. angiogenesis inhibitors for treating cancer and

PT diabetic retinopathy -

XX

PS Disclosure; Page 127; 146pp; English.

XX

CC The present invention provides endostatin peptides which can be used in  
CC the modulation of angiogenesis. This is useful in the treatment of  
CC cancers, inflammation, rheumatoid arthritis, chronic articular  
CC rheumatism, psoriasis, disorders associated with inopportune invasion of  
CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy  
CC of prematurity, macular degeneration, corneal graft rejection,  
CC retrolental fibroplasia, rubeosis, capillary proliferation in

CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent  
CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque  
CC neovascularisation, telangiectasia, haemophilic joints and wound  
CC granulation. In addition, the peptides can be used as birth control  
CC agents.

XX

SQ Sequence 7 AA;

Query Match 55.6%; Score 5; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRL 6

|||||

Db 1 LSSRL 5

#### RESULT 15

AAB49868

ID AAB49868 standard; Peptide; 9 AA.

XX

AC AAB49868;

XX

DT 02-MAR-2001 (first entry)

XX

DE Human endostatin peptide fragment SEQ ID NO: 85.

XX

KW Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken;  
KW cancer; inflammation; angiogenesis-dependent disease.

XX

OS Homo sapiens.

XX

PN WO200067771-A1.

XX

PD 16-NOV-2000.

XX

PF 02-MAY-2000; 2000WO-US12063.

XX

PR 06-MAY-1999; 99US-0132907.

PR 14-JUL-1999; 99US-0353333.

XX

PA (BURN-) BURNHAM INST.

XX

PI Vuori K;

XX

DR WPI; 2001-040937/05.

XX

PT Endostatin peptide comprising at least four endostatin amino acid  
PT residues are e.g. angiogenesis inhibitors for treating cancer and  
PT diabetic retinopathy -

XX

PS Disclosure; Page 139; 146pp; English.

XX

CC The present invention provides endostatin peptides which can be used in  
CC the modulation of angiogenesis. This is useful in the treatment of  
CC cancers, inflammation, rheumatoid arthritis, chronic articular  
CC rheumatism, psoriasis, disorders associated with inopportune invasion of

CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy  
CC of prematurity, macular degeneration, corneal graft rejection,  
CC retrolental fibroplasia, rubeosis, capillary proliferation in  
CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent  
CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque  
CC neovascularisation, telangiectasia, haemophilic joints and wound  
CC granulation. In addition, the peptides can be used as birth control  
CC agents.

XX

SQ Sequence 9 AA;

Query Match 55.6%; Score 5; DB 22; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRL 6  
|||||  
Db 1 LSSRL 5

Search completed: November 13, 2003, 10:32:52  
Job time : 28.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:38:27 ; Search time 17.625 Seconds  
(without alignments)  
93.222 Million cell updates/sec

Title: US-09-228-866-3  
Perfect score: 9  
Sequence: 1 CLSSRLDAC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 124183

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*

```

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	55.6	10	11	US-09-572-404B-23	Sequence 23, Appl
2	5	55.6	16	12	US-10-160-162-272	Sequence 272, App
3	5	55.6	16	12	US-09-820-649-272	Sequence 272, App
4	5	55.6	17	8	US-08-424-550B-241	Sequence 241, App
5	5	55.6	18	15	US-10-142-238A-84	Sequence 84, Appl
6	5	55.6	20	12	US-10-280-066-363	Sequence 363, App
7	4	44.4	7	9	US-09-989-789-1558	Sequence 1558, Ap
8	4	44.4	7	10	US-09-922-261-330	Sequence 330, App
9	4	44.4	7	11	US-09-990-186-1558	Sequence 1558, Ap
10	4	44.4	7	11	US-09-989-994-1558	Sequence 1558, Ap
11	4	44.4	8	12	US-10-190-082-370	Sequence 370, App
12	4	44.4	9	10	US-09-364-597A-19	Sequence 19, Appl
13	4	44.4	9	10	US-09-364-597A-20	Sequence 20, Appl
14	4	44.4	9	10	US-09-840-277-60	Sequence 60, Appl
15	4	44.4	9	10	US-09-840-277-61	Sequence 61, Appl
16	4	44.4	9	11	US-09-262-126C-6	Sequence 6, Appli
17	4	44.4	9	12	US-09-932-165-896	Sequence 896, App
18	4	44.4	9	12	US-09-932-165-1035	Sequence 1035, Ap
19	4	44.4	9	12	US-09-942-052-46	Sequence 46, Appl
20	4	44.4	9	12	US-09-942-052-113	Sequence 113, App
21	4	44.4	9	12	US-09-942-052-126	Sequence 126, App
22	4	44.4	9	12	US-09-942-052-216	Sequence 216, App
23	4	44.4	9	12	US-09-942-052-217	Sequence 217, App
24	4	44.4	9	12	US-09-942-052-306	Sequence 306, App
25	4	44.4	9	12	US-09-942-052-428	Sequence 428, App
26	4	44.4	9	12	US-10-210-148-91	Sequence 91, Appl
27	4	44.4	9	12	US-10-277-292-332	Sequence 332, App
28	4	44.4	9	12	US-10-280-340-332	Sequence 332, App
29	4	44.4	9	15	US-10-245-803-6	Sequence 6, Appli
30	4	44.4	10	9	US-09-834-765-289	Sequence 289, App
31	4	44.4	10	9	US-09-834-765-378	Sequence 378, App
32	4	44.4	10	9	US-09-796-848A-12	Sequence 12, Appl
33	4	44.4	10	10	US-09-996-288-22	Sequence 22, Appl
34	4	44.4	10	10	US-09-996-288-107	Sequence 107, App

35	4	44.4	10	11	US-09-572-404B-386	Sequence 386, App
36	4	44.4	10	11	US-09-572-404B-1781	Sequence 1781, Ap
37	4	44.4	10	11	US-09-996-265-22	Sequence 22, Appl
38	4	44.4	10	11	US-09-996-265-107	Sequence 107, App
39	4	44.4	10	12	US-09-932-165-1132	Sequence 1132, Ap
40	4	44.4	10	12	US-09-932-165-1338	Sequence 1338, Ap
41	4	44.4	10	12	US-09-572-270A-894	Sequence 894, App
42	4	44.4	10	12	US-09-942-052-59	Sequence 59, Appl
43	4	44.4	10	12	US-09-942-052-94	Sequence 94, Appl
44	4	44.4	10	12	US-09-942-052-195	Sequence 195, App
45	4	44.4	10	12	US-09-942-052-254	Sequence 254, App

# ALIGNMENTS

## RESULT 1

US-09-572-404B-23

```
; Sequence 23, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 23
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in CBFA2 OR AML1 at 289-298 and may
interact with
; OTHER INFORMATION: Sequence 24 in this patent.
US-09-572-404B-23
```

```
Query Match          55.6%; Score 5; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 53;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      2 LSSRL 6
        |||||
Db      2 LSSRL 6
```

## RESULT 2

US-10-160-162-272

```
; Sequence 272, Application US/10160162
; Publication No. US20030166541A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 83 Human Secreted Proteins
; FILE REFERENCE: PZ012P2
; CURRENT APPLICATION NUMBER: US/10/160,162
; CURRENT FILING DATE: 2002-06-04
```

```

; PRIOR APPLICATION NUMBER: 60/295,558
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/236,557
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: PCT/US98/15949
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,212
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,209
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,234
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,218
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,214
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,236
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,215
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,211
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,217
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/054,213
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/055,968
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,969
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,972
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,561
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,534
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,729
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,543
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,727
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,554
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,730
; PRIOR FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 272
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-160-162-272

```

Query Match 55.6%; Score 5; DB 12; Length 16;

Best Local Similarity 100.0%; Pred. No. 79;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRL 6  
|||||  
Db 7 LSSRL 11

RESULT 3

US-09-820-649-272

; Sequence 272, Application US/09820649  
; Publication No. US20030199683A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 83 Human Secreted Proteins  
; FILE REFERENCE: PZ012P1  
; CURRENT APPLICATION NUMBER: US/09/820,649  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US/09/236,557  
; PRIOR FILING DATE: 1999-01-26  
; PRIOR APPLICATION NUMBER: PCT/US98/15949  
; PRIOR FILING DATE: 1998-07-29  
; PRIOR APPLICATION NUMBER: 60/054,212  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,209  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,234  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,218  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,214  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,236  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,215  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/054,211  
; PRIOR FILING DATE: 1997-07-30  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 353  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 272  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-820-649-272

Query Match 55.6%; Score 5; DB 12; Length 16;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRL 6  
|||||  
Db 7 LSSRL 11

RESULT 4

US-08-424-550B-241

; Sequence 241, Application US/08424550B

```

; Publication No. US20020119447A1
; GENERAL INFORMATION:
;   APPLICANT: JOHN N. SIMONS
;   APPLICANT: TAMI J. PILOT-MATIAS
;   APPLICANT: GEORGE J. DAWSON
;   APPLICANT: GEORGE G. SCHLAUDER
;   APPLICANT: SURESH M. DESAI
;   APPLICANT: THOMAS P. LEARY
;   APPLICANT: ANTHONY SCOTT MUERHOFF
;   APPLICANT: JAMES C. ERKER
;   APPLICANT: SHERI L. BUIJK
;   APPLICANT: ISA K. MUSHAHWAR
;   TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;   TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;   NUMBER OF SEQUENCES: 716
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;     STREET: 100 ABBOTT PARK ROAD
;     CITY: ABBOTT PARK
;     STATE: IL
;     COUNTRY: USA
;     ZIP: 60064-3500
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/424,550B
;     FILING DATE:
;     CLASSIFICATION: 435435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: POREMBSKI, PRISCILLA E.
;     REGISTRATION NUMBER: 33,207
;     REFERENCE/DOCKET NUMBER: 5527.PC.01
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 708-937-6365
;     TELEFAX: 708-938-2623
;   INFORMATION FOR SEQ ID NO: 241:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 17 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-424-550B-241

```

```

Query Match          55.6%; Score 5; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 83;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      2 LSSRL 6
        |||||
Db      7 LSSRL 11

```

```

RESULT 5
US-10-142-238A-84

```



```

; Sequence 84, Application US/10142238A
; Publication No. US20030087819A1
; GENERAL INFORMATION:
; APPLICANT: Bielicki, John K.
; TITLE OF INVENTION: CYSTEINE-CONTAINING PEPTIDES HAVING OXIDANT PROPERTIES
; FILE REFERENCE: IB-1705
; CURRENT APPLICATION NUMBER: US/10/142,238A
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/289,944
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 84
; LENGTH: 18
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(18)
; OTHER INFORMATION: SYNTHETIC SEQUENCE
US-10-142-238A-84

```

```

Query Match          55.6%; Score 5; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 87;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 SRLDA 8
        |||||
Db      10 SRLDA 14

```

# RESULT 6

```

US-10-280-066-363
; Sequence 363, Application US/10280066
; Publication No. US20030180718A1
; GENERAL INFORMATION:
; APPLICANT: Pillutla, Renuka C.
; APPLICANT: Brissette, Renee
; APPLICANT: Spruyt, Michael
; APPLICANT: Dedova, Olga
; APPLICANT: Blume, Arthur J.
; APPLICANT: Prendergast, John
; APPLICANT: Goldstein, Neil I.
; TITLE OF INVENTION: TARGET SPECIFIC SCREENING AND ITS USE FOR IDENTIFYING
TARGET BINDERS
; FILE REFERENCE: 2598-4009US1
; CURRENT APPLICATION NUMBER: US/10/280,066
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/345,471
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 537
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 363
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Eschericia coli
; FEATURE:

```

; NAME/KEY: MISC\_FEATURE  
; OTHER INFORMATION: VEGF-20F-3-D7  
US-10-280-066-363

Query Match 55.6%; Score 5; DB 12; Length 20;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSRL 6  
|||||  
Db 1 LSSRL 5

RESULT 7

US-09-989-789-1558  
; Sequence 1558, Application US/09989789  
; Patent No. US20020063379A1  
; GENERAL INFORMATION:  
; APPLICANT: LIU, Qiang  
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE  
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS  
; FILE REFERENCE: 8325-0011.20 / S11-US2  
; CURRENT APPLICATION NUMBER: US/09/989,789  
; CURRENT FILING DATE: 2002-03-25  
; NUMBER OF SEQ ID NOS: 4085  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1558  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP  
US-09-989-789-1558

Query Match 44.4%; Score 4; DB 9; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
|||||  
Db 2 SSRL 5

RESULT 8

US-09-922-261-330  
; Sequence 330, Application US/09922261  
; Patent No. US20020111471A1  
; GENERAL INFORMATION:  
; APPLICANT: COGENT NEUROSCIENCE, Inc.  
; APPLICANT: Lo, Donald C.  
; APPLICANT: Barney, Shawn  
; APPLICANT: Thomas, Mary Beth  
; APPLICANT: Portbury, Stuart D.  
; APPLICANT: Puranam, Kasturi  
; APPLICANT: Katz, Lawrence C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING

; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING

; TITLE OF INVENTION: CELL DEATH  
; FILE REFERENCE: 10001-005-999  
; CURRENT APPLICATION NUMBER: US/09/922,261  
; CURRENT FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US/09/461,697  
; PRIOR FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 466  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 330  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-922-261-330

Query Match 44.4%; Score 4; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SRLD 7  
    ||||  
Db 4 SRLD 7

#### RESULT 9

US-09-990-186-1558  
; Sequence 1558, Application US/09990186  
; Publication No. US20030068675A1  
; GENERAL INFORMATION:  
; APPLICANT: LIU, Qiang  
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE  
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS  
; FILE REFERENCE: 8325-0011.21 / S11-US3  
; CURRENT APPLICATION NUMBER: US/09/990,186  
; CURRENT FILING DATE: 2001-11-20  
; NUMBER OF SEQ ID NOS: 4085  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1558  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP  
US-09-990-186-1558

Query Match 44.4%; Score 4; DB 11; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
    ||||  
Db 2 SSRL 5

#### RESULT 10

US-09-989-994-1558

; Sequence 1558, Application US/09989994  
; Publication No. US20030104526A1  
; GENERAL INFORMATION:  
; APPLICANT: LIU, Qiang  
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE  
; TITLE OF INVENTION: TRIPLETS BY ZINC FINGERS  
; FILE REFERENCE: 8325-0011.20 / S11-US2  
; CURRENT APPLICATION NUMBER: US/09/989,994  
; CURRENT FILING DATE: 2001-11-20  
; NUMBER OF SEQ ID NOS: 4085  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1558  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: example ZFP  
US-09-989-994-1558

Query Match 44.4%; Score 4; DB 11; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
|||  
Db 2 SSRL 5

RESULT 11

US-10-190-082-370

; Sequence 370, Application US/10190082  
; Publication No. US20030148264A1  
; GENERAL INFORMATION:  
; APPLICANT: Lasky, Lawrence A.  
; APPLICANT: Sidhu, Sachdev S.  
; APPLICANT: Held, Heike A.  
; TITLE OF INVENTION: PHAGE DISPLAYED PDZ DOMAIN LIGANDS  
; FILE REFERENCE: P1905R1  
; CURRENT APPLICATION NUMBER: US/10/190,082  
; CURRENT FILING DATE: 2002-07-03  
; PRIOR APPLICATION NUMBER: US 60/303,634  
; PRIOR FILING DATE: 2001-07-06  
; NUMBER OF SEQ ID NOS: 683  
; SEQ ID NO 370  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-190-082-370

Query Match 44.4%; Score 4; DB 12; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
|||

Db

5 SSRL 8

RESULT 12

US-09-364-597A-19

; Sequence 19, Application US/09364597A  
; Patent No. US20020103130A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruoslahti, Erkki  
; APPLICANT: Koivunen, Erkki  
; TITLE OF INVENTION: No. US20020103130A1e1 Integrin-Binding Peptides  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/364,597A  
; FILING DATE: 30-JUL-1999  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/158,001  
; FILING DATE: 24-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,861  
; FILING DATE: 04-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 3419  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (858) 535-9001  
; TELEFAX: (858) 535-8949  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: circular  
US-09-364-597A-19

Query Match 44.4%; Score 4; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RLDA 8  
|||  
Db 4 RLDA 7

RESULT 13

US-09-364-597A-20

```
; Sequence 20, Application US/09364597A
; Patent No. US20020103130A1
; GENERAL INFORMATION:
;   APPLICANT: Ruoslahti, Erkki
;   APPLICANT: Koivunen, Erkki
;   TITLE OF INVENTION: No. US20020103130A1e1 Integrin-Binding Peptides
;   NUMBER OF SEQUENCES: 46
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Campbell & Flores LLP
;     STREET: 4370 La Jolla Village Drive, Suite 700
;     CITY: San Diego
;     STATE: California
;     COUNTRY: USA
;     ZIP: 92122
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/364,597A
;     FILING DATE: 30-JUL-1999
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/158,001
;     FILING DATE: 24-NOV-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/286,861
;     FILING DATE: 04-AUG-1994
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Campbell, Cathryn
;     REGISTRATION NUMBER: 31,815
;     REFERENCE/DOCKET NUMBER: P-LA 3419
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (858) 535-9001
;     TELEFAX: (858) 535-8949
;   INFORMATION FOR SEQ ID NO: 20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 9 amino acids
;       TYPE: amino acid
;       TOPOLOGY: circular
```

US-09-364-597A-20

```
Query Match          44.4%; Score 4; DB 10; Length 9;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      4 SRLD 7
        ||||
Db      3 SRLD 6
```

RESULT 14

US-09-840-277-60

```
; Sequence 60, Application US/09840277
```

```

; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Integrin antagonist peptide
US-09-840-277-60

```

```

Query Match          44.4%;  Score 4;  DB 10;  Length 9;
Best Local Similarity 100.0%;  Pred. No. 6e+05;
Matches      4;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      5 RLDA 8
        ||||
Db      4 RLDA 7

```

# RESULT 15

```

US-09-840-277-61
; Sequence 61, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:

```

; OTHER INFORMATION: Integrin antagonist peptide  
US-09-840-277-61

Query Match 44.4%; Score 4; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SRLD 7  
|||  
Db 3 SRLD 6

Search completed: November 13, 2003, 11:12:32  
Job time : 17.625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:26:01 ; Search time 8.8125 Seconds  
(without alignments)  
98.215 Million cell updates/sec

Title: US-09-228-866-3  
Perfect score: 9  
Sequence: 1 CLSSRLDAC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3752

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

		%				
Result		Query				
No.	Score	Match	Length	DB	ID	Description



1	4	44.4	15	2	PH0770	T-cell receptor be
2	4	44.4	19	2	E56661	S-locus specific g
3	4	44.4	20	2	I67551	monocyte chemotact
4	3	33.3	10	2	B33143	pneumadin - human
5	3	33.3	10	2	A33143	pneumadin - rat
6	3	33.3	11	1	ECLQ2M	tachykinin II - mi
7	3	33.3	11	2	PT0217	T-cell receptor be
8	3	33.3	11	2	D41946	T-cell receptor ga
9	3	33.3	12	1	LFECPE	pyrE leader peptid
10	3	33.3	12	2	PN0577	tyrosine 3-monooxy
11	3	33.3	12	2	PN0578	tyrosine 3-monooxy
12	3	33.3	12	2	PN0579	tyrosine 3-monooxy
13	3	33.3	12	2	PN0580	tyrosine 3-monooxy
14	3	33.3	12	2	PN0581	tyrosine 3-monooxy
15	3	33.3	12	2	PN0576	tyrosine 3-monooxy
16	3	33.3	12	2	S26541	T-cell receptor be
17	3	33.3	12	2	S49547	hypothetical prote
18	3	33.3	12	2	PH0746	T-cell receptor be
19	3	33.3	13	2	PQ0491	self-incompatibili
20	3	33.3	13	2	D56661	S-locus specific g
21	3	33.3	13	2	S47359	T-cell antigen rec
22	3	33.3	13	2	S47377	T-cell antigen rec
23	3	33.3	13	2	S47384	T-cell antigen rec
24	3	33.3	13	2	S47388	T-cell antigen rec
25	3	33.3	13	2	B47415	mannose-1-phosphat
26	3	33.3	13	2	S01043	glutamate-ammonia
27	3	33.3	14	2	PH1471	T-cell receptor be
28	3	33.3	14	2	PH0945	T-cell receptor be
29	3	33.3	15	2	S49409	H+-transporting tw
30	3	33.3	15	2	B56661	S-locus specific g
31	3	33.3	15	2	PA0029	protein QA100012 -
32	3	33.3	15	2	G41299	T-cell receptor al
33	3	33.3	15	2	B49655	T-cell-receptor be
34	3	33.3	15	2	PH0772	T-cell receptor be
35	3	33.3	16	2	B25979	nodulation protein
36	3	33.3	16	2	S35627	uvsX protein - pha
37	3	33.3	16	2	PH0137	T-cell receptor be
38	3	33.3	16	2	A28587	T-cell receptor be
39	3	33.3	17	2	I55226	myosin heavy chain
40	3	33.3	17	2	I57941	beta 3-adrenergic
41	3	33.3	17	2	A29834	trp leader peptide
42	3	33.3	17	2	B86323	protein F14D16.5 [
43	3	33.3	17	2	JP0068	ribosomal protein
44	3	33.3	17	2	S50742	proteinase inhibit
45	3	33.3	17	2	E28587	T-cell receptor be

#### ALIGNMENTS

##### RESULT 1

PH0770

T-cell receptor beta chain (L4) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 05-Nov-1999

C;Accession: PH0770

R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.  
 J. Exp. Med. 174, 1371-1383, 1991  
 A;Title: T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: implications for T cell allelic exclusion and antigen-specific repertoire.  
 A;Reference number: PH0746; MUID:92078846; PMID:1836010  
 A;Accession: PH0770  
 A;Molecule type: mRNA  
 A;Residues: 1-15 <CAS>  
 A;Cross-references: EMBL:X60864; NID:g52855; PIDN:CAA43254.1; PID:g52856  
 A;Experimental source: T lymphocyte  
 C;Keywords: T-cell receptor

Query Match 44.4%; Score 4; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
 ||||  
 Db 3 SSRL 6

#### RESULT 2

E56661  
 S-locus specific glycoprotein (allele S2) - wild cabbage (fragment)  
 C;Species: Brassica oleracea (wild cabbage)  
 C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-Jan-1996  
 C;Accession: E56661  
 R;Gaude, T.; Denoroy, L.; Dumas, C.  
 Electrophoresis 12, 646-653, 1991  
 A;Title: Use of a fast protein electrophoretic purification procedure for N-terminal sequence analysis to identify S-locus related proteins in stigmas of Brassica oleracea.  
 A;Reference number: A56661; MUID:92090397; PMID:1752245  
 A;Accession: E56661  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-19 <GAU>  
 A;Experimental source: stigma extracts, var. alboglabra  
 A;Note: sequence extracted from NCBI backbone (NCBIP:72299)  
 C;Comment: This glycoprotein, expressed only in stigmas, plays an important role in the prevention of self-fertilization.  
 C;Keywords: glycoprotein; polymorphism

Query Match 44.4%; Score 4; DB 2; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSR 5  
 ||||  
 Db 4 LSSR 7

#### RESULT 3

I67551  
 monocyte chemotactic protein - human (fragment)

C;Species: Homo sapiens (man)  
 C;Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jul-2000  
 C;Accession: I67551  
 R;Steenbergen, E.J.; Verhagen, O.J.; van Leeuwen, E.F.; Behrendt, H.; Merle, P.A.; Wester, M.R.; von dem Borne, A.E.; van der Schoot, C.E.  
 Eur. J. Immunol. 24, 900-908, 1994  
 A;Title: B precursor acute lymphoblastic leukemia third complementarity-determining regions predominantly represent an unbiased recombination repertoire: leukemic transformation frequently occurs in fetal life.  
 A;Reference number: I53401; MUID:94200227; PMID:8149961  
 A;Accession: I67551  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-20 <RES>  
 A;Cross-references: GB:S69743; NID:g546304; PIDN:AAD14041.1; PID:g4261741  
 C;Genetics:  
 A;Gene: IgH VDJ

Query Match 44.4%; Score 4; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRLD 7  
 ||||  
 Db 16 SRLD 19

RESULT 4  
 B33143  
 pneumadin - human  
 C;Species: Homo sapiens (man)  
 C;Date: 16-Nov-1990 #sequence\_revision 16-Nov-1990 #text\_change 18-Aug-2000  
 C;Accession: B33143  
 R;Batra, V.K.; Mathur, M.; Mir, S.A.; Kapoor, R.; Kumar, M.A.  
 Regul. Pept. 30, 77-87, 1990  
 A;Title: Pneumadin: a new lung peptide which triggers antidiuresis.  
 A;Reference number: A33143; MUID:91110910; PMID:2274681  
 A;Accession: B33143  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-10 <BAT>  
 C;Superfamily: unassigned animal peptides

Query Match 33.3%; Score 3; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LDA 8  
 |||  
 Db 6 LDA 8

RESULT 5  
 A33143  
 pneumadin - rat  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 16-Nov-1990 #sequence\_revision 16-Nov-1990 #text\_change 18-Aug-2000

C;Accession: A33143  
R;Batra, V.K.; Mathur, M.; Mir, S.A.; Kapoor, R.; Kumar, M.A.  
Regul. Pept. 30, 77-87, 1990  
A;Title: Pneumadin: a new lung peptide which triggers antidiuresis.  
A;Reference number: A33143; MUID:91110910; PMID:2274681  
A;Accession: A33143  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-10 <BAT>  
C;Superfamily: unassigned animal peptides

Query Match 33.3%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
|||  
Db 6 LDA 8

#### RESULT 6

ECLQ2M

tachykinin II - migratory locust

C;Species: Locusta migratoria (migratory locust)

C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 08-Dec-1995

C;Accession: S08266

R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; de Loof, A.

FEBS Lett. 261, 397-401, 1990

A;Title: Locustatachykinin I and II, two novel insect neuropeptides with  
homology to peptides of the vertebrate tachykinin family.

A;Reference number: S08265; MUID:90184489; PMID:2311766

A;Accession: S08266

A;Molecule type: protein

A;Residues: 1-11 <SCH>

C;Superfamily: tachykinin

C;Keywords: amidated carboxyl end; neuropeptide; tachykinin

F;11/Modified site: amidated carboxyl end (Arg) #status experimental

Query Match 33.3%; Score 3; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSS 4  
|||  
Db 3 LSS 5

#### RESULT 7

PT0217

T-cell receptor beta chain V-J region (4-1-E.2) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 30-May-1997

C;Accession: PT0217

R;Nakano, N.; Kikutani, H.; Nishimoto, H.; Kishimoto, T.

J. Exp. Med. 173, 1091-1097, 1991

A;Title: T cell receptor V gene usage of islet beta cell-reactive T cells is not  
restricted in non-obese diabetic mice.

A;Reference number: PT0209; MUID:91217621; PMID:1902501

A;Accession: PT0217

A;Molecule type: mRNA

A;Residues: 1-11 <NAK>

C;Keywords: T-cell receptor

Query Match 33.3%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SRL 6  
|||  
Db 3 SRL 5

#### RESULT 8

D41946

T-cell receptor gamma chain (1a.4) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 07-May-1999

C;Accession: D41946

R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.

Mol. Cell. Biol. 11, 5902-5909, 1991

A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma genes in intestinal intraepithelial lymphocytes from murine athymic chimeras.

A;Reference number: A41946; MUID:92049316; PMID:1658619

A;Accession: D41946

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-11 <WHE>

C;Keywords: T-cell receptor

Query Match 33.3%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSS 4  
|||  
Db 7 LSS 9

#### RESULT 9

LFECPE

pyrE leader peptide - Escherichia coli

C;Species: Escherichia coli

C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 31-Mar-1993

C;Accession: A30400; A05110; Q00495

R;Poulsen, P.; Bonekamp, F.; Jensen, K.F.

EMBO J. 3, 1783-1790, 1984

A;Title: Structure of the Escherichia coli pyrE operon and control of pyrE expression by a UTP modulated intercistronic attenuation.

A;Reference number: A30400; MUID:85003588; PMID:6207018

A;Accession: A30400

A;Molecule type: DNA

A;Residues: 1-12 <POU1>

R;Poulsen, P.; Jensen, K.F.; Valentin-Hansen, P.; Carlsson, P.; Lundberg, L.G.  
 Eur. J. Biochem. 135, 223-229, 1983  
 A;Title: Nucleotide sequence of the Escherichia coli pyrE gene and of the DNA in  
 front of the protein-coding region.  
 A;Reference number: A05110; MUID:83287414; PMID:6349999  
 A;Accession: A05110  
 A;Molecule type: DNA  
 A;Residues: 1-12 <POU2>  
 C;Genetics:  
 A;Gene: pyrE-LP  
 A;Map position: 82 min  
 C;Superfamily: pyrE leader peptide

Query Match 33.3%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SRL 6  
 |||  
 Db 2 SRL 4

# RESULT 10

PN0577

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - gorilla (fragment)  
 N;Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine  
 hydroxylase

C;Species: Gorilla gorilla (gorilla)

C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 31-Mar-2000

C;Accession: PN0577

R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.  
 Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A;Reference number: PN0575; MUID:93371398; PMID:7689834

A;Accession: PN0577

A;Molecule type: genomic RNA

A;Residues: 1-12 <ICH>

A;Cross-references: GB:L14795

A;Experimental source: lymphocytes of peripheral blood

C;Comment: This enzyme catalyzes the first and rate-limiting step of  
 catecholamine biosynthesis.

C;Superfamily: phenylalanine 4-monooxygenase

C;Keywords: biopterin; monooxygenase; oxidoreductase

Query Match 33.3%; Score 3; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
 |||  
 Db 3 LDA 5

# RESULT 11

PN0578

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - orangutan (fragment)

N;Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase  
 C;Species: Pongo pygmaeus (orangutan)  
 C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 31-Mar-2000  
 C;Accession: PN0578  
 R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.  
 Biochem. Biophys. Res. Commun. 195, 158-165, 1993  
 A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.  
 A;Reference number: PN0575; MUID:93371398; PMID:7689834  
 A;Accession: PN0578  
 A;Molecule type: genomic RNA  
 A;Residues: 1-12 <ICH>  
 A;Cross-references: GB:L14798  
 A;Experimental source: lymphocytes of peripheral blood  
 C;Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis.  
 C;Superfamily: phenylalanine 4-monooxygenase  
 C;Keywords: bioppterin; monooxygenase; oxidoreductase

Query Match 33.3%; Score 3; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
 |||  
 Db 3 LDA 5

# RESULT 12

PN0579  
 tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - common gibbon (fragment)  
 N;Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase  
 C;Species: Hylobates lar (common gibbon, white-handed gibbon)  
 C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 31-Mar-2000  
 C;Accession: PN0579  
 R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.  
 Biochem. Biophys. Res. Commun. 195, 158-165, 1993  
 A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.  
 A;Reference number: PN0575; MUID:93371398; PMID:7689834  
 A;Accession: PN0579  
 A;Molecule type: genomic RNA  
 A;Residues: 1-12 <ICH>  
 A;Cross-references: GB:L14792  
 A;Experimental source: lymph nodes  
 C;Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis.  
 C;Superfamily: phenylalanine 4-monooxygenase  
 C;Keywords: bioppterin; monooxygenase; oxidoreductase

Query Match 33.3%; Score 3; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
 |||  
 Db 3 LDA 5

RESULT 13

PN0580

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - Japanese macaque (fragment)

N;Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C;Species: *Macaca fuscata* (Japanese macaque)

C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 31-Mar-2000

C;Accession: PN0580

R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T. Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A;Reference number: PN0575; MUID:93371398; PMID:7689834

A;Accession: PN0580

A;Molecule type: genomic RNA

A;Residues: 1-12 <ICH>

A;Cross-references: GB:L14804

A;Experimental source: kidney

C;Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis.

C;Superfamily: phenylalanine 4-monooxygenase

C;Keywords: biopterin; monooxygenase; oxidoreductase

Query Match 33.3%; Score 3; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
|||  
Db 3 LDA 5

RESULT 14

PN0581

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - common marmoset (fragment)

N;Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C;Species: *Callithrix jacchus* (common marmoset)

C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 31-Mar-2000

C;Accession: PN0581

R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T. Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A;Reference number: PN0575; MUID:93371398; PMID:7689834

A;Accession: PN0581

A;Molecule type: genomic RNA

A;Residues: 1-12 <ICH>

A;Cross-references: GB:L14803

A;Experimental source: kidney

C;Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis.

C;Superfamily: phenylalanine 4-monooxygenase

C;Keywords: biopterin; monooxygenase; oxidoreductase



Query Match 33.3%; Score 3; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
|||  
Db 3 LDA 5

RESULT 15

PN0576

tyrosine 3-monooxygenase (EC 1.14.16.2) type 1 and 2 - chimpanzee (fragment)

N;Alternate names: oxygen oxidoreductase; tetrahydropteridine; tyrosine hydroxylase

C;Species: Pan troglodytes (chimpanzee)

C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 31-Mar-2000

C;Accession: PN0576

R;Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, T.

Biochem. Biophys. Res. Commun. 195, 158-165, 1993

A;Title: Increased heterogeneity of tyrosine hydroxylase in humans.

A;Reference number: PN0575; MUID:93371398; PMID:7689834

A;Accession: PN0576

A;Molecule type: genomic RNA

A;Residues: 1-12 <ICH>

A;Cross-references: GB:L14789

A;Experimental source: lymphocytes of peripheral blood

C;Comment: This enzyme catalyzes the first and rate-limiting step of catecholamine biosynthesis.

C;Superfamily: phenylalanine 4-monooxygenase

C;Keywords: biopterin; monooxygenase; oxidoreductase

Query Match 33.3%; Score 3; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 2.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
|||  
Db 3 LDA 5

Search completed: November 13, 2003, 10:39:53

Job time : 9.8125 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:55:06 ; Search time 4.875 Seconds  
(without alignments)  
86.819 Million cell updates/sec

Title: US-09-228-866-3  
Perfect score: 9  
Sequence: 1 CLSSRLDAC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1246

Minimum DB seq length: 7

Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4	44.4	19	1	FIBB_LAMGL	P14473 lama glama
2	4	44.4	21	1	FIBB_ANTAM	P14465 antilocapra
3	3	33.3	10	1	PNEU_HUMAN	P22103 homo sapien
4	3	33.3	10	1	PNEU_RAT	P21996 rattus norv
5	3	33.3	12	1	FRE1_LITIN	P82021 litoria inf
6	3	33.3	12	1	YZPY_ECOLI	P17776 escherichia
7	3	33.3	15	1	TA1_TREBR	P34070 tremella br
8	3	33.3	16	1	UVSX_BPT6	Q06728 bacteriopha
9	3	33.3	17	1	LPW_CORGL	P06556 corynebacte
10	3	33.3	17	1	UP31_UPEIN	P82032 uperoleia i
11	3	33.3	17	1	UP32_UPEIN	P82033 uperoleia i
12	3	33.3	17	1	UP33_UPEIN	P82034 uperoleia i
13	3	33.3	18	1	NPA_BOVIN	P15506 bos taurus
14	3	33.3	18	1	YMDB_CHLAU	Q45827 chloroflexu
15	3	33.3	19	1	FIBB_TAPTE	P14539 tapirus ter
16	3	33.3	20	1	FIBB_SHEEP	P14470 ovis aries
17	3	33.3	20	1	ITRA_ALBJU	P24925 albizzia ju
18	3	33.3	20	1	JHBP_BOMMO	P81627 bombyx mori
19	3	33.3	21	1	FER_PYRWO	P81638 pyrococcus
20	3	33.3	21	1	FIBB_CEREL	P14468 cervus elap
21	3	33.3	21	1	FIBB_MUNMU	P14475 muntiacus m
22	3	33.3	21	1	FIBB_ODOHE	P14476 odocoileus
23	3	33.3	21	1	FIBB_RANTA	P14479 rangifer ta
24	3	33.3	21	1	NRLA_ACISP	P33036 acinetobact
25	2	22.2	7	1	CARP_MYTED	P10420 mytilus edu
26	2	22.2	7	1	CHOX_ALCSP	P16101 alcaligenes
27	2	22.2	7	1	UF03_MOUSE	P38641 mus musculu
28	2	22.2	8	1	ALL3_CYPDPO	P82154 cydia pomon
29	2	22.2	8	1	COXG_RAT	P80430 rattus norv
30	2	22.2	8	1	CPD1_ENTFA	P13269 enterococcu
31	2	22.2	8	1	FAR4_MACRS	P83277 macrobrachi
32	2	22.2	8	1	LCK2_LEUMA	P21141 leucophaea

33	2	22.2	8	1	LCK4_LEUMA	P21143	leucophaea
34	2	22.2	8	1	LCK5_LEUMA	P19987	leucophaea
35	2	22.2	8	1	LCK6_LEUMA	P19988	leucophaea
36	2	22.2	8	1	LCK7_LEUMA	P19989	leucophaea
37	2	22.2	8	1	LMT2_LOCFI	P22396	locusta mig
38	2	22.2	8	1	LPK_LEUMA	P13049	leucophaea
39	2	22.2	8	1	LPMS_STAEP	P23211	staphylococ
40	2	22.2	8	1	ORMY_ORCLI	P82455	orconectes
41	2	22.2	8	1	PPK2_PERAM	P82692	periplaneta
42	2	22.2	8	1	PPK3_PERAM	P82618	periplaneta
43	2	22.2	9	1	DNF1_LOCFI	P16339	locusta mig
44	2	22.2	9	1	DSIP_RABIT	P01158	oryctolagus
45	2	22.2	9	1	FAR3_PENMO	P83318	penaeus mon

# ALIGNMENTS

## RESULT 1

### FIBB\_LAMGL

ID FIBB\_LAMGL STANDARD; PRT; 19 AA.  
AC P14473;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Lama glama (Llama),  
OS Lama vicugna (Vicugna) (Vicugna vicugna), and  
OS Camelus dromedarius (Dromedary) (Arabian camel).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
OX NCBI\_TaxID=9844, 9843, 9838;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=L.glama;  
RA Blomback B., Blomback M., Grondahl N.J.;  
RT "Studies on fibrinopeptides from mammals.";  
RL Acta Chem. Scand. 19:1789-1791(1965).  
RN [2]  
RP SEQUENCE.  
RC SPECIES=C.dromedarius;  
RX MEDLINE=67209145; PubMed=6033721;  
RA Doolittle R.F., Schubert D., Schwartz S.A.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.  
RT Dromedary camel, mule deer, and cape buffalo.";  
RL Arch. Biochem. Biophys. 118:456-467(1967).  
RN [3]  
RP SEQUENCE.  
RC SPECIES=L.vicugna;  
RA Mross G.A., Doolittle R.F.;  
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";  
RL Arch. Biochem. Biophys. 122:674-684(1967).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation.  
 FT PEPTIDE 1 19 FIBRINOPEPTIDE B.  
 FT MOD\_RES 4 4 SULFATION.  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 2295 MW; E7EE6B6100568638 CRC64;

Query Match 44.4%; Score 4; DB 1; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RLDA 8  
 ||||  
 Db 15 RLDA 18

## RESULT 2

### FIBB\_ANTAM

ID FIBB\_ANTAM STANDARD; PRT; 21 AA.  
 AC P14465;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Antilocapra americana (Pronghorn).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Antilocapridae; Antilocapra.  
 OX NCBI\_TaxID=9891;  
 RN [1]  
 RP SEQUENCE.  
 RA Mross G.A., Doolittle R.F.;  
 RT "Amino acid sequence studies on artiodacty fibrinopeptides.";  
 RL Arch. Biochem. Biophys. 122:674-684(1967).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma; Sulfation; Pyrrolidone carboxylic acid.  
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 6 6 SULFATION.  
 FT NON\_TER 21 21

SQ SEQUENCE 21 AA; 2585 MW; FCE6183BE1F31627 CRC64;

Query Match 44.4%; Score 4; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 RLDA 8  
||||  
Db 17 RLDA 20

#### RESULT 3

##### PNEU\_HUMAN

ID PNEU\_HUMAN STANDARD; PRT; 10 AA.  
AC P22103;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Pneumadin (PNM).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Lung;  
RX MEDLINE=91110910; PubMed=2274681;  
RA Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;  
RT "Pneumadin: a new lung peptide which triggers antidiuresis.";  
RL Regul. Pept. 30:77-87(1990).  
CC -!- FUNCTION: ANTIDIURETIC PEPTIDE THAT TRIGGERS THE RELEASE OF ADH.  
DR PIR; B33143; B33143.  
DR GO; GO:0030103; P:vasopressin secretion; NAS.  
KW Amidation.  
FT MOD\_RES 10 10 AMIDATION.  
SQ SEQUENCE 10 AA; 956 MW; 640378DAA723276B CRC64;

Query Match 33.3%; Score 3; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
|||  
Db 6 LDA 8

#### RESULT 4

##### PNEU\_RAT

ID PNEU\_RAT STANDARD; PRT; 10 AA.  
AC P21996;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Pneumadin (PNM).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Lung;  
 RX MEDLINE=91110910; PubMed=2274681;  
 RA Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;  
 RT "Pneumadin: a new lung peptide which triggers antidiuresis."  
 RL Regul. Pept. 30:77-87(1990).  
 CC -!- FUNCTION: THIS ANTIDIURETIC PEPTIDE TRIGGERS THE RELEASE OF ADH.  
 DR PIR; A33143; A33143.  
 KW Amidation.  
 FT MOD\_RES 10 10 AMIDATION.  
 SQ SEQUENCE 10 AA; 1048 MW; 641D00DAA723276B CRC64;

Query Match 33.3%; Score 3; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
 |||  
 Db 6 LDA 8

#### RESULT 5

FRE1\_LITIN  
 ID FRE1\_LITIN STANDARD; PRT; 12 AA.  
 AC P82021;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Frenatin 1.  
 OS Litoria infrafrrenata (Giant tree frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;  
 OC Pelodryadinae; Litoria.  
 OX NCBI\_TaxID=61195;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Skin secretion;  
 RX MEDLINE=97368637; PubMed=9225251;  
 RA Raftery M.J., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;  
 RT "The structures of the frenatin peptides from the skin secretion of  
 the giant tree frog Litoria infrafrrenata."  
 RL J. Pept. Sci. 2:117-124(1996).  
 CC -!- FUNCTION: Wide spectrum antimicrobial peptide.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the skin parotoid and/or rostral  
 CC glands.  
 CC -!- MASS SPECTROMETRY: MW=1140; METHOD=FAB.  
 KW Amphibian defense peptide; Antibiotic; Amidation.  
 FT MOD\_RES 12 12 AMIDATION.  
 SQ SEQUENCE 12 AA; 1141 MW; C622550BC365B72D CRC64;

Query Match 33.3%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            6 LDA 8  
               |||  
 Db            3 LDA 5

RESULT 6

YZPY\_ECOLI

ID YZPY\_ECOLI            STANDARD;            PRT;    12 AA.  
 AC P17776;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DE Hypothetical pyrE leader peptide.  
 GN PYRL OR PYRE-LP.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85003588; PubMed=6207018;  
 RA Poulsen P., Bonekamp F., Jensen K.F.;  
 RT "Structure of the Escherichia coli pyrE operon and control of pyrE  
 RT expression by a UTP modulated intercistronic attenuation.";  
 RL EMBO J. 3:1783-1790(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=83287414; PubMed=6349999;  
 RA Poulsen P., Jensen K.F., Valentin-Hansen P., Carlsson P.,  
 RA Lundberg L.G.;  
 RT "Nucleotide sequence of the Escherichia coli pyrE gene and of the DNA  
 RT in front of the protein-coding region.";  
 RL Eur. J. Biochem. 135:223-229(1983).  
 CC -!- CAUTION: THIS SEQUENCE, ACCORDING TO THE ECOSEQ DATABASE (K. RUDD)  
 CC IS PROBABLY NOT A REAL PROTEIN; THEREFORE THIS ENTRY WILL PROBABLY  
 CC BE DELETED IN FUTURE RELEASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X00781; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; V01578; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A30400; LFECPE.  
 KW Hypothetical protein.  
 SQ SEQUENCE    12 AA;    1542 MW;    C4291FA437A2C9C9 CRC64;

Query Match                    33.3%;    Score 3;    DB 1;    Length 12;  
 Best Local Similarity    100.0%;    Pred. No. 8e+02;  
 Matches    3;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy            4 SRL 6

Db

|||  
2 SRL 4

RESULT 7

TA1\_TREBR

ID TA1\_TREBR STANDARD; PRT; 15 AA.  
AC P34070;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1994 (Rel. 28, Last annotation update)  
DE Tremmerogen A-I.  
OS Tremella brasiliensis (Jelly fungus).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellomycetidae; Tremellales; Tremellaceae; Tremella.  
OX NCBI\_TaxID=29896;  
RN [1]  
RP SEQUENCE.  
RA Ishibashi Y., Sakagami Y., Isogai A., Suzuki A.;  
RT "Structures of Tremmerogens A-9291-I and A-9291-VIII: peptidyl sex  
RT hormones of Tremella brasiliensis.";  
RL Biochemistry 23:1399-1404(1984).  
CC -!- FUNCTION: TREMEROGEN A-I IS PRODUCED BY THE A MATING-TYPE CELLS  
CC AND INDUCES FORMATION OF CONJUGATION TUBES IN A MATING-TYPE CELLS.  
KW Pheromone; Prenylation; Lipoprotein.  
FT LIPID 15 15 FARNESYL.  
SQ SEQUENCE 15 AA; 1339 MW; 3AABA4FC2D605333 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 9.9e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSR 5

|||  
Db 4 SSR 6

RESULT 8

UVSX\_BPT6

ID UVSX\_BPT6 STANDARD; PRT; 16 AA.  
AC Q06728;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Recombination and repair protein (Fragment).  
GN UVSX.  
OS Bacteriophage T6.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OX NCBI\_TaxID=10666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93219141; PubMed=8464751;  
RA Winkler M., Rueger W.;  
RT "Cloning and sequencing of the genes of  
RT beta-glucosyl-HMC-alpha-glucosyl-transferases of bacteriophages T2  
RT and T6.";



RL Nucleic Acids Res. 21:1500-1500(1993).  
 CC -!- FUNCTION: IMPORTANT IN GENETIC RECOMBINATION, DNA REPAIR, AND  
 CC REPLICATION. POSSESSES PAIRING AND STRAND-TRANSFER ACTIVITY.  
 CC INTERACTS WITH DDA AND GENE 32 PROTEINS.  
 CC -!- SIMILARITY: Belongs to the recA family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X68725; CAA48668.1; -.  
 DR PIR; S35627; S35627.  
 KW DNA damage; DNA replication; DNA recombination; DNA repair;  
 KW ATP-binding.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1721 MW; 48164C95D76F3CB4 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SRL 6  
 |||  
 Db 8 SRL 10

# RESULT 9

LPW\_CORGL  
 ID LPW\_CORGL STANDARD; PRT; 17 AA.  
 AC P06556;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Trp operon leader peptide.  
 GN TRPL OR CGL3028.1.  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87117512; PubMed=3808947;  
 RA Matsui K., Sano K., Ohtsubo E.;  
 RT "Complete nucleotide and deduced amino acid sequences of the  
 RT Brevibacterium lactofermentum tryptophan operon."  
 RL Nucleic Acids Res. 14:10113-10114(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88032866; PubMed=3667535;  
 RA Matsui K., Miwa K., Sano K.;  
 RT "Two single-base-pair substitutions causing desensitization to  
 RT tryptophan feedback inhibition of anthranilate synthase and enhanced  
 RT expression of tryptophan genes of Brevibacterium lactofermentum.";

RL J. Bacteriol. 169:5330-5332(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87277409; PubMed=3609747;  
 RA Sano K., Matsui K.;  
 RT "Structure and function of the trp operon control regions of  
 RT Brevibacterium lactofermentum, a glutamic-acid-producing bacterium.";  
 RL Gene 53:191-200(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13059 / AS019;  
 RX MEDLINE=91088299; PubMed=2263476;  
 RA Heery D.M., Dunican L.K.;  
 RT "Nucleotide sequence of the Corynebacterium glutamicum trpE gene.";  
 RL Nucleic Acids Res. 18:7138-7138(1990).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;  
 RA Nakagawa S.;  
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS  
 CC OF TRYPTOPHAN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X04960; CAA28622.1; -.  
 DR EMBL; M17892; AAB59110.1; -.  
 DR EMBL; M16663; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; X55994; CAA39466.1; -.  
 DR EMBL; AP005283; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A29834; A29834.  
 KW Tryptophan biosynthesis; Leader peptide; Complete proteome.  
 SQ SEQUENCE 17 AA; 2112 MW; 74C7E7924DAAE56B CRC64;

Query Match 33.3%; Score 3; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CLS 3  
 |||  
 Db 5 CLS 7

RESULT 10  
 UP31\_UPEIN  
 ID UP31\_UPEIN STANDARD; PRT; 17 AA.  
 AC P82032;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Uperin 3.1.  
 OS Uperoleia inundata (Floodplain toadlet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
 OC Myobatrachinae; Uperoleia.  
 OX NCBI\_TaxID=104953;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Skin secretion;  
 RA Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,  
 RA Adams G.W., Severini C.;  
 RT "Novel uperin peptides from the dorsal glands of the australian  
 RT floodplain toadlet Uperoleia inundata.";  
 RL Aust. J. Chem. 49:475-484(1996).  
 CC -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST S.AUREUS  
 CC AND L.MESENTERIOIDES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.  
 CC -!- MASS SPECTROMETRY: MW=1827; METHOD=FAB.  
 KW Amphibian defense peptide; Antibiotic; Amidation.  
 FT MOD\_RES 17 17 AMIDATION.  
 SQ SEQUENCE 17 AA; 1829 MW; 7E5B748FAC83DB93 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LDA 8  
 |||  
 Db 3 LDA 5

# RESULT 11

## UP32\_UPEIN

ID UP32\_UPEIN STANDARD; PRT; 17 AA.  
 AC P82033;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Uperin 3.2.  
 OS Uperoleia inundata (Floodplain toadlet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
 OC Myobatrachinae; Uperoleia.  
 OX NCBI\_TaxID=104953;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Skin secretion;  
 RA Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,  
 RA Adams G.W., Severini C.;  
 RT "Novel uperin peptides from the dorsal glands of the australian  
 RT floodplain toadlet Uperoleia inundata.";  
 RL Aust. J. Chem. 49:475-484(1996).  
 CC -!- FUNCTION: UNKNOWN.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.  
 CC -!- MASS SPECTROMETRY: MW=1841; METHOD=FAB.

KW Amphibian defense peptide; Amidation.  
FT MOD\_RES 17 17 AMIDATION.  
SQ SEQUENCE 17 AA; 1843 MW; 7E4E948FAC83DB93 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LDA 8  
|||  
Db 3 LDA 5

RESULT 12

UP33\_UPEIN

ID UP33\_UPEIN STANDARD; PRT; 17 AA.  
AC P82034;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Uperin 3.3.  
OS Uperoleia inundata (Floodplain toadlet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
OC Myobatrachinae; Uperoleia.  
OX NCBI\_TaxID=104953;  
RN [1]  
RP SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=Skin secretion;  
RA Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,  
RA Adams G.W., Severini C.;  
RT "Novel uperin peptides from the dorsal glands of the australian  
RT floodplain toadlet Uperoleia inundata.";  
RL Aust. J. Chem. 49:475-484(1996).  
CC -!- FUNCTION: UNKNOWN.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.  
CC -!- MASS SPECTROMETRY: MW=1813; METHOD=FAB.  
KW Amphibian defense peptide; Amidation.  
FT MOD\_RES 17 17 AMIDATION.  
SQ SEQUENCE 17 AA; 1815 MW; 7E4E8BB2AC83DB93 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LDA 8  
|||  
Db 3 LDA 5

RESULT 13

NPA\_BOVIN

ID NPA\_BOVIN STANDARD; PRT; 18 AA.  
AC P15506;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neuropeptide A.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=86067985; PubMed=3865193;  
 RA Yang H.-Y.T., Fratta W., Majane E.A., Costa E.;  
 RT "Isolation, sequencing, synthesis, and pharmacological  
 RT characterization of two brain neuropeptides that modulate the action  
 RT of morphine."  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7757-7761(1985).  
 CC -!- FUNCTION: MODULATES THE ACTION OF MORPHINE.  
 DR PIR; A24749; A24749.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 18 18 AMIDATION.  
 SQ SEQUENCE 18 AA; 1921 MW; EC52DAE1F45CFCFB CRC64;

Query Match 33.3%; Score 3; DB 1; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSS 4  
 |||  
 Db 5 LSS 7

#### RESULT 14 YMDB\_CHLAU

ID YMDB\_CHLAU STANDARD; PRT; 18 AA.  
 AC Q45827;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein in mdh 3'region (ORFB) (Fragment).  
 OS Chloroflexus aurantiacus.  
 OC Bacteria; Chloroflexi; Chloroflexales; Chloroflexaceae; Chloroflexus.  
 OX NCBI\_TaxID=1108;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J-10-fl;  
 RX MEDLINE=96241868; PubMed=8661927;  
 RA Synstad B., Emmerhoff O., Sirevag R.;  
 RT "Malate dehydrogenase from the green gliding bacterium Chloroflexus  
 RT aurantiacus is phylogenetically related to lactic dehydrogenases."  
 RL Arch. Microbiol. 165:346-353(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC -----  
DR EMBL; X89038; CAA61437.1; -.  
KW Hypothetical protein.  
FT NON\_TER 18 18  
SQ SEQUENCE 18 AA; 1797 MW; 1A09256C224075EB CRC64;

Query Match 33.3%; Score 3; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSS 4  
|||  
Db 13 LSS 15

#### RESULT 15

##### FIBB\_TAPTE

ID FIBB\_TAPTE STANDARD; PRT; 19 AA.  
AC P14539;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Tapirus terrestris (Lowland tapir) (Brazilian tapir).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.  
OX NCBI\_TaxID=9801;  
RN [1]  
RP SEQUENCE.  
RA O'Neil P.B., Doolittle R.F.;  
RT "Mammalian phylogeny based on fibrinopeptide amino acid sequences.";  
RL Syst. Zool. 22:590-595(1973).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Sulfation.  
FT PEPTIDE 1 19 FIBRINOPEPTIDE B.  
FT MOD\_RES 4 4 SULFATION.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2226 MW; FCEA265C35E99627 CRC64;

Query Match 33.3%; Score 3; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDA 8  
|||

Db 16 LDA 18

Search completed: November 13, 2003, 10:33:57  
Job time : 4.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:58:36 ; Search time 22.125 Seconds  
(without alignments)  
104.971 Million cell updates/sec

Title: US-09-228-866-3  
Perfect score: 9  
Sequence: 1 CLSSRLDAC 9

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4	44.4	9	10	Q9FSZ2	Q9fsz2 cicer ariet
2	4	44.4	11	2	P77404	P77404 escherichia
3	4	44.4	17	10	O65345	O65345 gossypium h
4	4	44.4	19	4	Q96ER8	Q96er8 homo sapien
5	4	44.4	19	5	Q8T8B6	Q8t8b6 ciona intes
6	4	44.4	19	10	Q9S921	Q9s921 brassica ol
7	4	44.4	20	5	Q27385	Q27385 trypanosoma
8	4	44.4	21	4	Q96D65	Q96d65 homo sapien
9	3	33.3	7	13	O42564	O42564 fugu rubrip
10	3	33.3	8	2	Q51594	Q51594 escherichia
11	3	33.3	9	2	P72149	P72149 pseudomonas
12	3	33.3	9	4	Q9UC36	Q9uc36 homo sapien
13	3	33.3	10	2	Q9L5W6	Q9l5w6 liberibacte
14	3	33.3	10	13	P82080	P82080 limnodynast
15	3	33.3	10	15	Q8UT83	Q8ut83 human immun
16	3	33.3	11	8	Q33551	Q33551 crithidia f
17	3	33.3	12	4	Q9HBU4	Q9hbu4 homo sapien
18	3	33.3	12	15	Q8AEW7	Q8aew7 human immun
19	3	33.3	13	2	Q9AMK2	Q9amk2 xanthomonas
20	3	33.3	13	2	Q9AIR1	Q9air1 pseudomonas
21	3	33.3	13	2	O86260	O86260 klebsiella
22	3	33.3	13	3	O75014	O75014 schizosacch
23	3	33.3	13	4	Q15537	Q15537 homo sapien
24	3	33.3	13	6	Q9TRW6	Q9trw6 bos taurus
25	3	33.3	13	10	Q9S922	Q9s922 brassica ol
26	3	33.3	14	2	P83159	P83159 anabaena sp
27	3	33.3	14	10	Q9S939	Q9s939 beta vulgar
28	3	33.3	14	12	Q9W982	Q9w982 hepatitis g
29	3	33.3	15	2	P83000	P83000 pseudomonas
30	3	33.3	15	3	Q96VH6	Q96vh6 penicillium
31	3	33.3	15	4	Q9UC67	Q9uc67 homo sapien
32	3	33.3	15	4	P78533	P78533 homo sapien
33	3	33.3	15	4	Q9UBM5	Q9ubm5 homo sapien
34	3	33.3	15	8	Q8SL36	Q8sl36 monanthes i
35	3	33.3	15	8	Q8HIF8	Q8hif8 arthroderma
36	3	33.3	15	8	Q8HIF6	Q8hif6 trichophyto
37	3	33.3	15	10	Q9AT15	Q9at15 lycopersico
38	3	33.3	15	10	Q9S924	Q9s924 brassica ol
39	3	33.3	15	10	Q9SAP4	Q9sap4 solanum tub
40	3	33.3	15	11	Q05694	Q05694 mus musculu
41	3	33.3	15	15	Q80362	Q80362 human immun
42	3	33.3	16	4	Q9UP51	Q9up51 homo sapien
43	3	33.3	16	4	Q96RT5	Q96rt5 homo sapien
44	3	33.3	16	4	Q8J026	Q8j026 homo sapien
45	3	33.3	16	6	Q9TR09	Q9tr09 bos taurus

## ALIGNMENTS

RESULT 1  
Q9FSZ2



ID Q9FSZ2 PRELIMINARY; PRT; 9 AA.  
AC Q9FSZ2;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Hypothetical 1.0 kDa protein (Fragment).  
OS Cicer arietinum (Chickpea) (Garbanzo).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.  
OX NCBI\_TaxID=3827;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Castellana; TISSUE=Etiolated epicotyl;  
RA Dopico B., Jimenez T., Labrador E.;  
RT "cDNA clones expressed in etiolated Cicer arietinum epicotyls.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AJ299069; CAC10216.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 990 MW; 9441BDDAA7272EBE CRC64;

Query Match 44.4%; Score 4; DB 10; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LDAC 9  
| | | |  
Db 5 LDAC 8

## RESULT 2

P77404

ID P77404 PRELIMINARY; PRT; 11 AA.  
AC P77404;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE DNA sequence downstream of the ECOPRR1 HSD locus (Fragment).  
GN HSDR.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97206151; PubMed=9157244;  
RA Tyndall C., Lehnerr H., Sandmeier U., Kulik E., Bickle T.A.;  
RT "The type IC hsd loci of the enterobacteria are flanked by DNA with  
RT high homology to the phage P1 genome: implications for the evolution  
RT and spread of DNA restriction systems.";  
RL Mol. Microbiol. 23:729-736(1997).  
DR EMBL; X98145; CAA66840.1; -.  
DR EMBL; X98144; CAA66839.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 11 AA; 1259 MW; 714AB092A4072734 CRC64;

Query Match 44.4%; Score 4; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
    ||||  
Db 2 SSRL 5

RESULT 3

O65345

ID O65345 PRELIMINARY; PRT; 17 AA.  
AC O65345;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)  
DE Vacuolar proton ATPase 16 kDa proteolipid subunit (Fragment).  
GN VATPASE.  
OS Gossypium hirsutum (Upland cotton).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
OX NCBI\_TaxID=3635;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yamamoto E., Baird W.V.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF064202; AAC16556.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 17 AA; 1770 MW; 32A8B533D1A2A410 CRC64;

Query Match 44.4%; Score 4; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSR 5  
    ||||  
Db 7 LSSR 10

RESULT 4

Q96ER8

ID Q96ER8 PRELIMINARY; PRT; 19 AA.  
AC Q96ER8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis, and Embryonic carcinoma;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; BC011998; AAH11998.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 19 AA; 2180 MW; 137DB66BEE0B9B59 CRC64;

Query Match 44.4%; Score 4; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
|||  
Db 10 SSRL 13

#### RESULT 5

Q8T8B6

ID Q8T8B6 PRELIMINARY; PRT; 19 AA.  
AC Q8T8B6;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Choline acetyltransferase (Fragment).  
GN CI-CHAT.  
OS Ciona intestinalis.  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
OC Phlebobranchia; Cionidae; Ciona.  
OX NCBI\_TaxID=7719;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21863848; PubMed=11875658;  
RA Takamura K., Egawa T., Ohnishi S., Okada T., Fukuoka T.;  
RT "Developmental expression of ascidian neurotransmitter synthesis  
RT genesI. Choline acetyltransferase and acetylcholine transporter  
RT genes.";  
RL Dev. Genes Evol. 212:50-53(2002).  
DR EMBL; AB072000; BAB85861.1; -.  
KW Transferase.  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 2176 MW; 36DC3BCDC6AB922F CRC64;

Query Match 44.4%; Score 4; DB 5; Length 19;  
Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
|||  
Db 14 SSRL 17

#### RESULT 6

Q9S921

ID Q9S921 PRELIMINARY; PRT; 19 AA.  
AC Q9S921;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE S2 SGP=S-locus specific glycoprotein (Fragment).

OS Brassica oleracea (Cauliflower).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3712;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=92090397; PubMed=1752245;  
 RA Gaude T., Denoroy L., Dumas C.;  
 RT "Use of a fast protein electrophoretic purification procedure for N-  
 RT terminal sequence analysis to identify S-locus related proteins in  
 RT stigmas of Brassica oleracea.";  
 RL Electrophoresis 12:646-653(1991).  
 FT NON\_TER 1 1  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 2094 MW; 2C7D5E558C266C31 CRC64;

Query Match 44.4%; Score 4; DB 10; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSSR 5  
 ||||  
 Db 4 LSSR 7

# RESULT 7

Q27385

ID Q27385 PRELIMINARY; PRT; 20 AA.  
 AC Q27385;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
 DE Variable surface glycoprotein basic COPY (BC) (Fragment).  
 OS Trypanosoma brucei brucei.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_TaxID=5702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84242548; PubMed=6737319;  
 RA Murphy W.J., Brentano S.T., Rice-Ficht A.C., Dorfman D.M.,  
 RA Donelson J.E.;  
 RT "DNA rearrangements of the variable surface antigen genes of the  
 RT trypanosomes.";  
 RL J. Protozool. 31:65-73(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84212810; PubMed=6327735;  
 RA Donelson J.E., Murphy W.J., Brentano S.T., Rice-Ficht A.C., Cain G.D.;  
 RT "Comparison of the expression-linked extra copy (ELC) and basic copy  
 RT (BC) genes of a trypanosome surface antigen.";  
 RL J. Cell. Biochem. 23:1-12(1983).  
 DR EMBL; M31526; AAA30301.1; -.  
 DR EMBL; M29387; AAA30243.1; -.  
 DR EMBL; M29476; AAA30240.1; -.  
 DR EMBL; M29478; AAA30242.1; -.  
 DR EMBL; M30740; AAA30309.1; -.

DR EMBL; M30742; AAA30314.1; -.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2212 MW; 94956D70759F5209 CRC64;

Query Match 44.4%; Score 4; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CLSS 4  
|||  
Db 5 CLSS 8

RESULT 8

Q96D65

ID Q96D65 PRELIMINARY; PRT; 21 AA.  
AC Q96D65;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; BC012895; AAH12895.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 21 AA; 2355 MW; 3F3C4E923AAB9327 CRC64;

Query Match 44.4%; Score 4; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSRL 6  
|||  
Db 18 SSRL 21

RESULT 9

O42564

ID O42564 PRELIMINARY; PRT; 7 AA.  
AC O42564;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Truncated voltage-gated sodium channel alpha subunit (Fragment).  
GN SCN8A.  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetradontoidea; Tetraodontidae; Takifugu.

OX NCBI\_TaxID=31033;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97442476; PubMed=9295353;  
 RA Plummer N.W., McBurney M.W., Meisler M.H.;  
 RT "Alternative splicing of the sodium channel SCN8A predicts a truncated  
 RT two-domain protein in fetal brain and non-neuronal cells.";  
 RL J. Biol. Chem. 272:24008-24015(1997).  
 DR EMBL; U97673; AAB80916.1; -.  
 KW Ionic channel.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 7 AA; 730 MW; 75B72EA2C73772A0 CRC64;

Query Match 33.3%; Score 3; DB 13; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLS 3  
 |||  
 Db 5 CLS 7

# RESULT 10

Q51594

ID Q51594 PRELIMINARY; PRT; 8 AA.  
 AC Q51594;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CopB protein (Fragment).  
 OS Escherichia coli.  
 OG Plasmid ColV2-K94.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86223772; PubMed=2423502;  
 RA Weber P.C., Palchaudhuri S.;  
 RT "Incompatibility repressor in a repA-like replicon of the IncFI  
 RT plasmid ColV2-K94.";  
 RL J. Bacteriol. 166:1106-1112(1986).  
 DR EMBL; M13472; AAA23194.1; -.  
 KW Plasmid.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 8 AA; 998 MW; 0F37205AA73416D7 CRC64;

Query Match 33.3%; Score 3; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RLD 7  
 |||  
 Db 3 RLD 5

# RESULT 11

P72149

ID P72149 PRELIMINARY; PRT; 9 AA.  
AC P72149;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Putative glucokinase (Fragment).  
GN GLK.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=96427344; PubMed=8830708;  
RA Sage A.E., Proctor W.D., Phibbs P.V.Jr.;  
RT "A two-component response regulator, gltR, is required for glucose  
RT transport activity in Pseudomonas aeruginosa PA01.";  
RL J. Bacteriol. 178:6064-6066(1996).  
DR EMBL; U50932; AAC44474.1; -.  
KW Kinase.  
FT NON\_TER 1 1  
SQ SEQUENCE 9 AA; 944 MW; C3071DDAA72DC6C6 CRC64;

Query Match 33.3%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LDA 8  
|||  
Db 5 LDA 7

RESULT 12

Q9UC36

ID Q9UC36 PRELIMINARY; PRT; 9 AA.  
AC Q9UC36;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE 28 kDa heat shock protein homolog fragment 1 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92218434; PubMed=1560006;  
RA Kato K., Shinohara H., Goto S., Inaguma Y., Morishita R., Asano T.;  
RT "Copurification of small heat shock protein with alpha B crystallin  
RT from human skeletal muscle.";  
RL J. Biol. Chem. 267:7718-7725(1992).  
FT NON\_TER 1 1  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1220 MW; 26933415B1F77B43 CRC64;

Query Match 33.3%; Score 3; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SRL 6  
|||  
Db 6 SRL 8

RESULT 13

Q9L5W6

ID Q9L5W6 PRELIMINARY; PRT; 10 AA.  
AC Q9L5W6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE RplA (Fragment).  
GN RPLA.  
OS *Liberibacter africanus* subsp. *capensis*.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Candidatus *Liberibacter*.  
OX NCBI\_TaxID=119494;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Garnier M., Jagoueix-Eveillard S., Cronje P., Le Roux H., Bove J.;  
RT "Genomic characterization of a *Liberibacter* present in an ornamental  
RT rutaceous tree, *Calodendrum capense*, in the western Cape province of  
RT South Africa. Proposal of 'candidatus *Liberibacter africanus*  
RT subspecies *capensis*'.";  
RL Int. J. Syst. Evol. Microbiol. 0:0-0(2000).  
DR EMBL; AF248498; AAF68451.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 10 AA; 1094 MW; C42894D5A5B72AB2 CRC64;

Query Match 33.3%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSS 4  
|||  
Db 5 LSS 7

RESULT 14

P82080

ID P82080 PRELIMINARY; PRT; 10 AA.  
AC P82080;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE DYNASTIN 2.  
OS *Limnodynastes dumerilii* (Eastern Banjo Frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
OC Limnodynastinae; Limnodynastes.  
OX NCBI\_TaxID=104065;  
RN [1]



RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=TIBIAL GLAND;  
 RA Raftery M.J., Bradford A.M., Bowie J.H., Wallace J.C., Tyler M.J.;  
 RT "Peptides from Australian frogs. The structure of the dynastins from  
 RT the banjo frogs Limnodynastes interioras, Limnodynastes dumerilii and  
 RT Limnodynastes terraereginae.";  
 RL Aust. J. Chem. 46:833-842(1993).  
 CC -!- MASS SPECTROMETRY: MW=986; METHOD=FAB.  
 KW Amphibian skin.  
 SQ SEQUENCE 10 AA; 986 MW; 550FDC287735A5B7 CRC64;

Query Match 33.3%; Score 3; DB 13; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSS 4  
 |||  
 Db 3 LSS 5

# RESULT 15

Q8UT83

ID Q8UT83 PRELIMINARY; PRT; 10 AA.  
 AC Q8UT83;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Vpu protein.  
 GN VPU.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=00BW1795.6;  
 RA Novitsky V.A., Smith U.R., Gilbert P., McLane M.F., Chigwedere P.,  
 RA Williamson C., Ndung'u T., Klein I., Chang S.-Y., Peter T., Thior I.,  
 RA Foley B.T., Gaolekwe S., Rybak N., Gaseitsiwe S., Vannberg F.,  
 RA Marlink R., Lee T.-H., Essex M.;  
 RT "HIV-1 subtype C molecular phylogeny: consensus sequence for an AIDS  
 RT vaccine design.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF443097; AAL34766.1; -.  
 SQ SEQUENCE 10 AA; 1264 MW; 91E52CB33321B37A CRC64;

Query Match 33.3%; Score 3; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SRL 6  
 |||  
 Db 8 SRL 10

Search completed: November 13, 2003, 10:38:09  
 Job time : 23.125 secs

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:27:36 ; Search time 9 Seconds  
(without alignments)  
37.610 Million cell updates/sec

Title: US-09-228-866-4  
Perfect score: 8  
Sequence: 1 CVLRGGRC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118358

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	8	100.0	8	1	US-08-526-710-4	Sequence 4, Appli
2	8	100.0	8	3	US-08-862-855-4	Sequence 4, Appli
3	8	100.0	8	3	US-09-226-985-4	Sequence 4, Appli
4	8	100.0	8	4	US-09-227-906-4	Sequence 4, Appli
5	5	62.5	16	4	US-09-128-344A-61	Sequence 61, Appl
6	5	62.5	16	4	US-09-128-344A-63	Sequence 63, Appl
7	5	62.5	16	4	US-09-128-344A-186	Sequence 186, App
8	5	62.5	16	4	US-09-128-344A-188	Sequence 188, App
9	5	62.5	17	4	US-09-128-344A-13	Sequence 13, Appl
10	5	62.5	17	4	US-09-128-344A-21	Sequence 21, Appl
11	5	62.5	17	4	US-09-128-344A-47	Sequence 47, Appl

12	5	62.5	17	4	US-09-128-344A-93	Sequence 93, Appl
13	5	62.5	17	4	US-09-128-344A-95	Sequence 95, Appl
14	5	62.5	17	4	US-09-128-344A-105	Sequence 105, App
15	5	62.5	17	4	US-09-128-344A-109	Sequence 109, App
16	5	62.5	17	4	US-09-128-344A-122	Sequence 122, App
17	5	62.5	17	4	US-09-128-344A-126	Sequence 126, App
18	5	62.5	17	4	US-09-128-344A-130	Sequence 130, App
19	5	62.5	17	4	US-09-128-344A-143	Sequence 143, App
20	5	62.5	18	2	US-08-752-852A-15	Sequence 15, Appl
21	5	62.5	18	4	US-09-128-344A-1	Sequence 1, Appli
22	5	62.5	18	4	US-09-128-344A-2	Sequence 2, Appli
23	5	62.5	18	4	US-09-128-344A-4	Sequence 4, Appli
24	5	62.5	18	4	US-09-128-344A-5	Sequence 5, Appli
25	5	62.5	18	4	US-09-128-344A-11	Sequence 11, Appl
26	5	62.5	18	4	US-09-128-344A-17	Sequence 17, Appl
27	5	62.5	18	4	US-09-128-344A-19	Sequence 19, Appl
28	5	62.5	18	4	US-09-128-344A-29	Sequence 29, Appl
29	5	62.5	18	4	US-09-128-344A-31	Sequence 31, Appl
30	5	62.5	18	4	US-09-128-344A-33	Sequence 33, Appl
31	5	62.5	18	4	US-09-128-344A-45	Sequence 45, Appl
32	5	62.5	18	4	US-09-128-344A-51	Sequence 51, Appl
33	5	62.5	18	4	US-09-128-344A-60	Sequence 60, Appl
34	5	62.5	18	4	US-09-128-344A-68	Sequence 68, Appl
35	5	62.5	18	4	US-09-128-344A-69	Sequence 69, Appl
36	5	62.5	18	4	US-09-128-344A-70	Sequence 70, Appl
37	5	62.5	18	4	US-09-128-344A-76	Sequence 76, Appl
38	5	62.5	18	4	US-09-128-344A-86	Sequence 86, Appl
39	5	62.5	18	4	US-09-128-344A-88	Sequence 88, Appl
40	5	62.5	18	4	US-09-128-344A-104	Sequence 104, App
41	5	62.5	18	4	US-09-128-344A-107	Sequence 107, App
42	5	62.5	18	4	US-09-128-344A-108	Sequence 108, App
43	5	62.5	18	4	US-09-128-344A-113	Sequence 113, App
44	5	62.5	18	4	US-09-128-344A-114	Sequence 114, App
45	5	62.5	18	4	US-09-128-344A-115	Sequence 115, App

#### ALIGNMENTS

#### RESULT 1

US-08-526-710-4

; Sequence 4, Application US/08526710

; Patent No. 5622699

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Pasqualini, Renata

; TITLE OF INVENTION: Method of Identifying Molecules That

; TITLE OF INVENTION: Home to a Selected Organ In Vivo

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-526-710-4

```

```

Query Match          100.0%; Score 8; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

QY      1 CVLRGGRC 8
        |||||
Db      1 CVLRGGRC 8

```

# RESULT 2

```

US-08-862-855-4
; Sequence 4, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-4

```

```

Query Match          100.0%; Score 8; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative    0; Mismatches    0; Indels      0; Gaps      0;

```

```

Qy      1 CVLRGGRC 8
        |||||
Db      1 CVLRGGRC 8

```

```

RESULT 3
US-09-226-985-4
; Sequence 4, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-4

```

```

Query Match          100.0%; Score 8; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CVLRGGRC 8
        |||||
Db      1 CVLRGGRC 8

```

#### RESULT 4

US-09-227-906-4

```

; Sequence 4, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-4

```

```

Query Match          100.0%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CVLRGGRC 8
        |||||
Db      1 CVLRGGRC 8

```

```

RESULT 5
US-09-128-344A-61
; Sequence 61, Application US/09128344A
; Patent No. 6307016
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/128,344A
; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-128-344A-61

Query Match 62.5%; Score 5; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||  
Db 1 RGGRC 5

RESULT 6  
US-09-128-344A-63  
; Sequence 63, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/128,344A  
; CURRENT FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-128-344A-63

Query Match 62.5%; Score 5; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||  
Db 1 RGGRC 5

RESULT 7  
US-09-128-344A-186  
; Sequence 186, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.



```
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/128,344A
; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: DISULFID
; LOCATION: (5)...(16)
US-09-128-344A-186
```

```
Query Match          62.5%; Score 5; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 19;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      4 RGGRC 8
        |||||
Db      1 RGGRC 5
```

# RESULT 8

```
US-09-128-344A-188
; Sequence 188, Application US/09128344A
; Patent No. 6307016
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/128,344A
; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 188
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: DISULFID
; LOCATION: (5)...(16)
```

US-09-128-344A-188

Query Match 62.5%; Score 5; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 9

US-09-128-344A-13

; Sequence 13, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/128,344A  
; CURRENT FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-128-344A-13

Query Match 62.5%; Score 5; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 10

US-09-128-344A-21

; Sequence 21, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS

```
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/128,344A
; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-128-344A-21
```

```
Query Match          62.5%; Score 5; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 20;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      4 RGGRC 8
        |||||
Db      1 RGGRC 5
```

# RESULT 11

US-09-128-344A-47

```
; Sequence 47, Application US/09128344A
; Patent No. 6307016
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/128,344A
; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: MOD_RES
; LOCATION: (1)...(17)
; OTHER INFORMATION: All genetically encoded amino acids are in the
; OTHER INFORMATION: D-configuration
US-09-128-344A-47
```

Query Match 62.5%; Score 5; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 12

US-09-128-344A-93  
; Sequence 93, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/128,344A  
; CURRENT FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 93  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-128-344A-93

Query Match 62.5%; Score 5; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 13

US-09-128-344A-95  
; Sequence 95, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/128,344A

```

; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-128-344A-95

```

```

Query Match          62.5%; Score 5; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 20;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 RGGRC 8
        |||||
Db      1 RGGRC 5

```

#### RESULT 14

```

US-09-128-344A-105
; Sequence 105, Application US/09128344A
; Patent No. 6307016
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/128,344A
; CURRENT FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: DISULFID
; LOCATION: (5)...(16)
; NAME/KEY: DISULFID
; LOCATION: (8)...(13)
US-09-128-344A-105

```

```

Query Match          62.5%; Score 5; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 20;

```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 15

US-09-128-344A-109  
; Sequence 109, Application US/09128344A  
; Patent No. 6307016  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/128,344A  
; CURRENT FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 109  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
; NAME/KEY: DISULFID  
; LOCATION: (5)...(16)  
; NAME/KEY: DISULFID  
; LOCATION: (8)...(13)  
US-09-128-344A-109

Query Match 62.5%; Score 5; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

Search completed: November 13, 2003, 10:41:54  
Job time : 9 secs

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:53:11 ; Search time 25.6667 Seconds  
(without alignments)  
49.473 Million cell updates/sec

Title: US-09-228-866-4  
Perfect score: 8  
Sequence: 1 CVLRGGRC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 358712

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result	% Query		Match	Length	DB	ID	Description
No.	Score						
1	8	100.0	8	18	AAW13414	Brain homing pepti	
2	8	100.0	8	21	AAB07390	Brain homing pepti	
3	8	100.0	8	22	AAE11796	Phage peptide #4 t	
4	8	100.0	8	23	AAU10707	Brain homing pepti	
5	5	62.5	8	23	AAO17769	Human beta-defensi	
6	5	62.5	9	23	AAU71320	Human MHC class I	
7	5	62.5	9	23	AAU71433	Human MHC molecule	
8	5	62.5	10	23	AAU71377	Human MHC class I	
9	5	62.5	10	23	AAU71478	Human MHC molecule	
10	5	62.5	17	18	AAW23735	Tachytegrin and/or	
11	5	62.5	17	18	AAW23737	Tachytegrin and/or	
12	5	62.5	17	23	AAO17771	Human beta-defensi	
13	5	62.5	17	23	AAO17780	Human beta-defensi	
14	5	62.5	18	18	AAW36214	Antimicrobial prot	
15	5	62.5	18	18	AAW23726	Tachytegrin and/or	
16	5	62.5	18	18	AAW23756	Tachytegrin and/or	
17	5	62.5	18	18	AAW23748	Tachytegrin and/or	
18	5	62.5	18	18	AAW23750	Tachytegrin and/or	
19	5	62.5	18	18	AAW23752	Tachytegrin and/or	
20	5	62.5	18	18	AAW23728	Tachytegrin and/or	
21	5	62.5	18	18	AAW23760	Tachytegrin and/or	
22	5	62.5	18	18	AAW23730	Tachytegrin and/or	
23	5	62.5	18	18	AAW23758	Tachytegrin and/or	
24	5	62.5	18	20	AAW99404	Protegrin derivati	
25	4	50.0	7	19	AAW77187	Pharmaceutically a	
26	4	50.0	7	19	AAW48491	Integrin receptor	
27	4	50.0	7	20	AAW86182	Peptide used in a	
28	4	50.0	7	23	ABP62082	Human secreted pro	
29	4	50.0	7	23	ABG77669	Targetting peptide	
30	4	50.0	7	23	ABG77670	Targetting peptide	
31	4	50.0	7	23	AAE13259	Fusion peptide #3	
32	4	50.0	8	17	AAW05711	Minimal motif #8.	
33	4	50.0	8	19	AAW79116	Gly-ala polymer of	
34	4	50.0	8	20	AAAY40757	S5 derivative #10,	
35	4	50.0	8	20	AAAY16805	Heat shock protein	
36	4	50.0	8	20	AAW74083	Fragment of gastro	
37	4	50.0	8	21	AAB30096	Scaffold protein S	
38	4	50.0	8	23	ABJ10461	Recombinase domain	
39	4	50.0	8	23	ABG60560	Selective targetin	
40	4	50.0	9	15	AAR62148	HIV-1 gp120/41 pro	
41	4	50.0	9	17	AAW49327	Human leucocyte an	
42	4	50.0	9	20	AAAY48986	Membrane dipeptida	
43	4	50.0	9	22	AAU03764	Cyclic peptide inh	
44	4	50.0	9	22	AAB75916	Hepatitis B virus	
45	4	50.0	9	23	ABG60847	Hyalauronan (HA) b	

# ALIGNMENTS



RESULT 1

AAW13414

ID AAW13414 standard; Peptide; 8 AA.

XX

AC AAW13414;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;  
KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 1997-202359/18.

XX

PT Obtaining compound that homes to selected organ or tissue - by in  
PT vivo panning method, specifically to identify brain, kidney,  
PT angiogenic vasculature or tumour tissue homing peptide(s)

XX

PS Claim 13; Page 67; 75pp; English.

XX

CC This synthetic peptide is a claimed example of a brain-homing  
CC peptide that was identified using a novel method for obtaining  
CC molecules that home to a selected organ or tissue. This in vivo  
CC panning method typically involves administering a phage display  
CC library to a subject, and identifying expressed peptides which  
CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
CC vascular tissue or tumour tissue. The isolated peptides (see  
CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
CC labels to the selected organ/tissue (claimed) or to identify and/or  
CC isolate target molecules (claimed). The peptides can be directly  
CC identified in vivo, as compared to prior art in vitro screening  
CC methods, which require further examination to see if they maintain  
CC specificity in vivo.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 18; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1 CVLRGGRC 8

Db

|||||||  
1 CVLRGGRC 8

RESULT 2

AAB07390

ID AAB07390 standard; peptide; 8 AA.

XX

AC AAB07390;

XX

DT 17-OCT-2000 (first entry)

XX

DE Brain homing peptide # 4.

XX

KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS Mus sp.

XX

FH Key Location/Qualifiers

FT Disulfide-bond 1..8

FT /note= "Can optionally form a cyclic peptide"

XX

PN US6068829-A.

XX

PD 30-MAY-2000.

XX

PF 23-JUN-1997; 97US-0862855.

XX

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 2000-410850/35.

XX

PT Identifying and recovering organ homing molecules or peptides by in

PT vivo panning comprises administering a library of diverse peptides

PT linked to a tag which facilitates recovery of these peptides -

XX

PS Example 2; Column 17; 20pp; English.

XX

CC The present sequence is a mouse brain homing peptide. This sequence was

CC identified by using in vivo panning to screen a library of potential

CC organ homing molecules. The present sequence can be used to direct a

CC moiety to a the brain tissue, by linking the moiety to the present

CC sequence. Examples of potential moieties are drugs, toxins or a

CC detectable label. The present sequence contains a VRL amino acid motif.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 21; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 CVLRGGRC 8

Db

|||||||  
1 CVLRGGRC 8

RESULT 3

AAE11796

ID AAE11796 standard; peptide; 8 AA.

XX

AC AAE11796;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #4 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;

KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

FH Key Location/Qualifiers

FT Domain 2..4

FT /label= VLR\_motif

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo

PT panning that selectively home to a selected organ or tissue useful for

PT treating disease or in diagnostic methods -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 22; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CVLRGGRC 8  
|||  
Db 1 CVLRGGRC 8

RESULT 4

AAU10707

ID AAU10707 standard; peptide; 8 AA.

XX

AC AAU10707;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #4 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
PT identifying molecules that home to a specific organ or tissue, e.g.  
PT identifying a tumour homing molecule to identify the presence of cancer,  
PT by in vivo panning of a library -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
CC home to a selected organ or tissue. The method comprises administering  
CC to the subject the library of diverse molecules, collecting a sample of  
CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
CC the sample several molecules that home to the selected organ or tissue.  
CC The method is useful for identifying molecules, particularly useful for  
CC screening large number of molecules (e.g. peptides), that home to a  
CC specific organ. The identified molecule is useful for e.g. raising an  
CC antibody specific for a target molecule, targeting a desired moiety

CC (e.g. drug, toxin or detectable label) to the selected organ.  
CC Specifically, the method is useful for identifying the presence of cancer  
CC in a subject by linking an appropriate moiety to a tumour homing  
CC molecule. The present method provides a direct means for identifying  
CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 23; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVLRGGRC 8  
|||  
Db 1 CVLRGGRC 8

#### RESULT 5

AAO17769

ID AAO17769 standard; peptide; 8 AA.

XX

AC AAO17769;

XX

DT 30-AUG-2002 (first entry)

XX

DE Human beta-defensin-3 derivative #4.

XX

KW Human; beta-defensin-3; hBD-3; bacterial infection; gene therapy;  
KW respiratory system; cystic fibrosis; inflammation; urogenital tract;  
KW antibacterial; fungicide; cytostatic; antiinflammatory; antiulcer;  
KW gastrointestinal tract; septicaemia; apoptosis induction; cancer.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "may be bound to between 1 and 50 amino acids"

FT Modified-site 8

FT /note= "may be bound to between 1 and 50 amino acids"

XX

PN WO200240512-A2.

XX

PD 23-MAY-2002.

XX

PF 14-NOV-2001; 2001WO-EP13174.

XX

PR 14-NOV-2000; 2000DE-1056365.

PR 30-MAR-2001; 2001DE-1016220.

XX

PA (IPFP-) IPF PHARM GMBH.

XX

PI Forssmann W, Kluever E, Conejo-Garcia J, Adermann K, Bals R;

PI Maegert H;  
 XX  
 DR WPI; 2002-435959/46.  
 XX  
 PT New human beta-defensin 3, useful for treating or preventing microbial  
 PT infection and tumors, also related nucleic acid -  
 XX  
 PS Claim 3; Page 24; 36pp; German.  
 XX  
 CC The present invention relates to human beta-defensin-3 (hBD-3) and its  
 CC derivatives. The peptide, its coding sequence and vectors containing the  
 CC coding sequence are useful in (gene) therapy and diagnosis, especially  
 CC for preventing or treating a wide range of microbial infections  
 CC (particularly Burkholderia cepacia and Pseudomonas aeruginosa in the  
 CC respiratory tract, especially in cases of cystic fibrosis, and  
 CC Helicobacter pylori, also inflammatory diseases of the gastrointestinal  
 CC and urogenital tracts, sepsis and yeast infections), and for inducing  
 CC apoptosis for treating malignant melanoma and tumours. The present  
 CC sequence is a derivative of human BD-3.  
 XX  
 SQ Sequence 8 AA;  
  
 Query Match 62.5%; Score 5; DB 23; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
 |||||  
 Db 4 RGGRC 8

RESULT 6  
 AAU71320  
 ID AAU71320 standard; Peptide; 9 AA.  
 XX  
 AC AAU71320;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE Human MHC class I molecule HLA-A3 binding 103P3E8 peptide #27.  
 XX  
 KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;  
 KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;  
 KW single chain monoclonal antibody; serum; blood; urine; tissue; human;  
 KW chromosome 9q13-q21.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179557-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 12-APR-2001; 2001WO-US12181.  
 XX  
 PR 12-APR-2000; 2000US-196647P.  
 XX  
 PA (UROG-) UROGENESYS INC.

XX  
 PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;  
 PI Jakobovits A;  
 XX  
 DR WPI; 2002-061976/08.  
 XX  
 PT Monitoring 103P3E8 gene products in sample from patient (suspected of)  
 PT having cancer, useful for diagnosing, managing or treating cancers,  
 PT e.g. prostate cancer, comprises determining presence of aberrant  
 PT 103P3E8 gene products -  
 XX  
 PS Disclosure; Page 87; 128pp; English.  
 XX  
 CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and  
 CC peptide fragments of the protein. 103P3E8 exhibits tissue specific  
 CC expression in normal adult tissue, but it is also aberrantly expressed in  
 CC many cancers including tumours of the prostate, bladder, kidney, colon,  
 CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related  
 CC protein and peptide fragments and specific PCR primers are therefore  
 CC useful for diagnosing and treating cancer. A vector comprising a  
 CC polynucleotide which encodes a single chain monoclonal antibody, that  
 CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme  
 CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,  
 CC are both useful in the preparation of a composition for treating a  
 CC patient with a cancer that expresses 103P3E8. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P3E8 gene products in  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.  
 XX  
 SQ Sequence 9 AA;

Query Match 62.5%; Score 5; DB 23; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGGR 7  
 |||||  
 Db 5 LRGGR 9

# RESULT 7

AAU71433

ID AAU71433 standard; Peptide; 9 AA.

XX

AC AAU71433;

XX

DT 26-FEB-2002 (first entry)

XX

DE Human MHC molecule HLA-A11 binding 103P3E8 peptide #40.

XX

KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;  
 KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;  
 KW single chain monoclonal antibody; serum; blood; urine; tissue; human;  
 KW chromosome 9q13-q21.

XX

OS Homo sapiens.

XX

PN WO200179557-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 12-APR-2001; 2001WO-US12181.  
 XX  
 PR 12-APR-2000; 2000US-196647P.  
 XX  
 PA (UROG-) UROGENESYS INC.  
 XX  
 PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;  
 PI Jakobovits A;  
 XX  
 DR WPI; 2002-061976/08.  
 XX  
 PT Monitoring 103P3E8 gene products in sample from patient (suspected of)  
 PT having cancer, useful for diagnosing, managing or treating cancers,  
 PT e.g. prostate cancer, comprises determining presence of aberrant  
 PT 103P3E8 gene products -  
 XX  
 PS Disclosure; Page 90; 128pp; English.  
 XX  
 CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and  
 CC peptide fragments of the protein. 103P3E8 exhibits tissue specific  
 CC expression in normal adult tissue, but it is also aberrantly expressed in  
 CC many cancers including tumours of the prostate, bladder, kidney, colon,  
 CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related  
 CC protein and peptide fragments and specific PCR primers are therefore  
 CC useful for diagnosing and treating cancer. A vector comprising a  
 CC polynucleotide which encodes a single chain monoclonal antibody, that  
 CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme  
 CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,  
 CC are both useful in the preparation of a composition for treating a  
 CC patient with a cancer that expresses 103P3E8. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P3E8 gene products in  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.  
 XX  
 SQ Sequence 9 AA;

Query Match 62.5%; Score 5; DB 23; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGGR 7  
 |||||  
 Db 5 LRGGR 9

RESULT 8  
 AAU71377  
 ID AAU71377 standard; Peptide; 10 AA.  
 XX  
 AC AAU71377;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX



DE Human MHC class I molecule HLA-A3 binding 103P3E8 peptide #84.  
 XX  
 KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;  
 KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;  
 KW single chain monoclonal antibody; serum; blood; urine; tissue; human;  
 KW chromosome 9q13-q21.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179557-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 12-APR-2001; 2001WO-US12181.  
 XX  
 PR 12-APR-2000; 2000US-196647P.  
 XX  
 PA (UROG-) UROGENESYS INC.  
 XX  
 PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;  
 PI Jakobovits A;  
 XX  
 DR WPI; 2002-061976/08.  
 XX  
 PT Monitoring 103P3E8 gene products in sample from patient (suspected of)  
 PT having cancer, useful for diagnosing, managing or treating cancers,  
 PT e.g. prostate cancer, comprises determining presence of aberrant  
 PT 103P3E8 gene products -  
 XX  
 PS Disclosure; Page 89; 128pp; English.  
 XX  
 CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and  
 CC peptide fragments of the protein. 103P3E8 exhibits tissue specific  
 CC expression in normal adult tissue, but it is also aberrantly expressed in  
 CC many cancers including tumours of the prostate, bladder, kidney, colon,  
 CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related  
 CC protein and peptide fragments and specific PCR primers are therefore  
 CC useful for diagnosing and treating cancer. A vector comprising a  
 CC polynucleotide which encodes a single chain monoclonal antibody, that  
 CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme  
 CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,  
 CC are both useful in the preparation of a composition for treating a  
 CC patient with a cancer that expresses 103P3E8. The sequences can be used  
 CC in diagnostic methods to monitor the level of 103P3E8 gene products in  
 CC serum, blood, urine and tissue and to thereby detect the presence of  
 CC cancerous cells.  
 XX  
 SQ Sequence 10 AA;

Query Match 62.5%; Score 5; DB 23; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGGR 7  
 |||||  
 Db 5 LRGGR 9

RESULT 9

AAU71478

ID AAU71478 standard; Peptide; 10 AA.

XX

AC AAU71478;

XX

DT 26-FEB-2002 (first entry)

XX

DE Human MHC molecule HLA-A11 binding 103P3E8 peptide #85.

XX

KW 103P3E8; prostate; bladder; kidney; colon; lung; breast; rectum; stomach;

KW tumour; cancer; cytostatic; gene therapy; antibody therapy; ribozyme;

KW single chain monoclonal antibody; serum; blood; urine; tissue; human;

KW chromosome 9q13-q21.

XX

OS Homo sapiens.

XX

PN WO200179557-A2.

XX

PD 25-OCT-2001.

XX

PF 12-APR-2001; 2001WO-US12181.

XX

PR 12-APR-2000; 2000US-196647P.

XX

PA (UROG-) UROGENESYS INC.

XX

PI Faris M, Challita-Eid PM, Raitano AB, Mitchell SC, Afar DEH;

PI Jakobovits A;

XX

DR WPI; 2002-061976/08.

XX

PT Monitoring 103P3E8 gene products in sample from patient (suspected of)

PT having cancer, useful for diagnosing, managing or treating cancers,

PT e.g. prostate cancer, comprises determining presence of aberrant

PT 103P3E8 gene products -

XX

PS Disclosure; Page 91; 128pp; English.

XX

CC Sequences AAU71093-AAU71796 represent the 103P3E8-related protein and

CC peptide fragments of the protein. 103P3E8 exhibits tissue specific

CC expression in normal adult tissue, but it is also aberrantly expressed in

CC many cancers including tumours of the prostate, bladder, kidney, colon,

CC lung, breast, rectum and stomach. The 103P3E8 polynucleotide, its related

CC protein and peptide fragments and specific PCR primers are therefore

CC useful for diagnosing and treating cancer. A vector comprising a

CC polynucleotide which encodes a single chain monoclonal antibody, that

CC immunospecifically binds to an 103P3E8-related protein, and a ribozyme

CC capable of cleaving a polynucleotide having the 103P3E8 coding sequence,

CC are both useful in the preparation of a composition for treating a

CC patient with a cancer that expresses 103P3E8. The sequences can be used

CC in diagnostic methods to monitor the level of 103P3E8 gene products in

CC serum, blood, urine and tissue and to thereby detect the presence of

CC cancerous cells.

XX

SQ Sequence 10 AA;

Query Match 62.5%; Score 5; DB 23; Length 10;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGGR 7  
      |||||  
Db 6 LRGGR 10

RESULT 10

AAW23735

ID AAW23735 standard; peptide; 17 AA.

XX

AC AAW23735;

XX

DT 16-OCT-1997 (first entry)

XX

DE Tachytegrin and/or parevin compound peptide.

XX

KW Protegrin; tachyplesin; antimicrobial; antiviral; oral mucositis;  
KW sexually transmitted disease; endotoxin; HIV; Staphylococcus aureus;  
KW Pseudomonas; H. pylori; Chlamydia trachomatis; Treponema pallidum;  
KW Neisseria gonorrhoeae; Trichomonas vaginalis; Herpes simplex type 2;  
KW animal; plant.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Disulfide-bond 5..17

FT /label= Optional

FT Disulfide-bond 8..13

FT /label= Optional

FT Modified-site 17

FT /note= "Optionally amidated"

XX

PN WO9702287-A1.

XX

PD 23-JAN-1997.

XX

PF 03-JUL-1996; 96WO-US11323.

XX

PR 06-JUL-1995; 95US-0000898.

XX

PA (INTR-) INTRABIOTICS PHARM INC.

XX

PI Chang CC, Gu CL, Harwig SSL, Lehrer RI;

XX

DR WPI; 1997-108914/10.

XX

PT Cpds. based on protegrin(s) and tachyplesin(s) - designated  
PT tachytegrin(s) and parevin(s), useful as antimicrobial and antiviral  
PT agents, e.g. for treatment of sexually transmitted diseases

XX

PS Claim 21; Page 43; 59pp; English.

XX

CC The present sequence represents a peptide sequence found in compounds

CC based on protegrins and tachyplesins, designated tachytegrins and  
 CC parevins. Both compounds can be used as preservatives (e.g. in  
 CC foodstuffs, cosmetics or medicaments) or in treatment of viral or  
 CC microbial infection in animals or plants. They are useful, e.g. in  
 CC treatment of oral mucositis, infection by Staphylococcus aureus,  
 CC Pseudomonas or H. pylori, or sexually transmitted diseases, including  
 CC those caused by Chlamydia trachomatis, Treponema pallidum, Neisseria  
 CC gonorrhoeae, Trichomonas vaginalis, Herpes simplex type 2 and HIV. DNA  
 CC encoding the compounds can also be expressed in situ, in animals or  
 CC plants, to combat infections. The compounds can also be used as  
 CC standards in antimicrobial assays and in binding endotoxins.

XX

SQ Sequence 17 AA;

Query Match 62.5%; Score 5; DB 18; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
 |||||  
 Db 1 RGGRC 5

# RESULT 11

AAW23737

ID AAW23737 standard; peptide; 17 AA.

XX

AC AAW23737;

XX

DT 16-OCT-1997 (first entry)

XX

DE Tachytegrin and/or parevin compound peptide.

XX

KW Protegrin; tachyplesin; antimicrobial; antiviral; oral mucositis;  
 KW sexually transmitted disease; endotoxin; HIV; Staphylococcus aureus;  
 KW Pseudomonas; H. pylori; Chlamydia trachomatis; Treponema pallidum;  
 KW Neisseria gonorrhoeae; Trichomonas vaginalis; Herpes simplex type 2;  
 KW animal; plant.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Disulfide-bond 5..15

FT /label= Optional

FT Disulfide-bond 8..13

FT /label= Optional

FT Modified-site 17

FT /note= "Optionally amidated"

XX

PN W09702287-A1.

XX

PD 23-JAN-1997.

XX

PF 03-JUL-1996; 96WO-US11323.

XX

PR 06-JUL-1995; 95US-0000898.

XX

PA (INTR-) INTRABIOTICS PHARM INC.  
 XX  
 PI Chang CC, Gu CL, Harwig SSL, Lehrer RI;  
 XX  
 DR WPI; 1997-108914/10.  
 XX  
 PT Cpd. based on protegrin(s) and tachyplesin(s) - designated  
 PT tachytegrin(s) and parevin(s), useful as antimicrobial and antiviral  
 PT agents, e.g. for treatment of sexually transmitted diseases  
 XX  
 PS Claim 21; Page 43; 59pp; English.  
 XX  
 CC The present sequence represents a peptide sequence found in compounds  
 CC based on protegrins and tachyplesins, designated tachytegrins and  
 CC parevins. Both compounds can be used as preservatives (e.g. in  
 CC foodstuffs, cosmetics or medicaments) or in treatment of viral or  
 CC microbial infection in animals or plants. They are useful, e.g. in  
 CC treatment of oral mucositis, infection by Staphylococcus aureus,  
 CC Pseudomonas or H. pylori, or sexually transmitted diseases, including  
 CC those caused by Chlamydia trachomatis, Treponema pallidum, Neisseria  
 CC gonorrhoeae, Trichomonas vaginalis, Herpes simplex type 2 and HIV. DNA  
 CC encoding the compounds can also be expressed in situ, in animals or  
 CC plants, to combat infections. The compounds can also be used as  
 CC standards in antimicrobial assays and in binding endotoxins.  
 XX  
 SQ Sequence 17 AA;

Query Match 62.5%; Score 5; DB 18; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
 |||||  
 Db 1 RGGRC 5

RESULT 12  
 AAO17771  
 ID AAO17771 standard; peptide; 17 AA.  
 XX  
 AC AAO17771;  
 XX  
 DT 30-AUG-2002 (first entry)  
 XX  
 DE Human beta-defensin-3 fragment #1.  
 XX  
 KW Human; beta-defensin-3; hBD-3; bacterial infection; gene therapy;  
 KW respiratory system; cystic fibrosis; inflammation; urogenital tract;  
 KW antibacterial; fungicide; cytostatic; antiinflammatory; antiulcer;  
 KW gastrointestinal tract; septicaemia; apoptosis induction; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200240512-A2.  
 XX  
 PD 23-MAY-2002.  
 XX

PF 14-NOV-2001; 2001WO-EP13174.  
 XX  
 PR 14-NOV-2000; 2000DE-1056365.  
 PR 30-MAR-2001; 2001DE-1016220.  
 XX  
 PA (IPFP-) IPF PHARM GMBH.  
 XX  
 PI Forssmann W, Kluever E, Conejo-Garcia J, Adermann K, Bals R;  
 PI Maegert H;  
 XX  
 DR WPI; 2002-435959/46.  
 XX  
 PT New human beta-defensin 3, useful for treating or preventing microbial  
 PT infection and tumors, also related nucleic acid -  
 XX  
 PS Claim 4; Page 24; 36pp; German.  
 XX  
 CC The present invention relates to human beta-defensin-3 (hBD-3) and its  
 CC derivatives. The peptide, its coding sequence and vectors containing the  
 CC coding sequence are useful in (gene) therapy and diagnosis, especially  
 CC for preventing or treating a wide range of microbial infections  
 CC (particularly Burkholderia cepacia and Pseudomonas aeruginosa in the  
 CC respiratory tract, especially in cases of cystic fibrosis, and  
 CC Helicobacter pylori, also inflammatory diseases of the gastrointestinal  
 CC and urogenital tracts, sepsis and yeast infections), and for inducing  
 CC apoptosis for treating malignant melanoma and tumours. The present  
 CC sequence is a fragment of human BD-3.  
 XX  
 SQ Sequence 17 AA;

Query Match 62.5%; Score 5; DB 23; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
 |||||  
 Db 9 RGGRC 13

# RESULT 13

AAO17780

ID AAO17780 standard; peptide; 17 AA.

XX

AC AAO17780;

XX

DT 30-AUG-2002 (first entry)

XX

DE Human beta-defensin-3 fragment #9.

XX

KW Human; beta-defensin-3; hBD-3; bacterial infection; gene therapy;  
 KW respiratory system; cystic fibrosis; inflammation; urogenital tract;  
 KW antibacterial; fungicide; cytostatic; antiinflammatory; antiulcer;  
 KW gastrointestinal tract; septicaemia; apoptosis induction; cancer.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Cross-links 6..13  
 XX  
 PN WO200240512-A2.  
 XX  
 PD 23-MAY-2002.  
 XX  
 PF 14-NOV-2001; 2001WO-EP13174.  
 XX  
 PR 14-NOV-2000; 2000DE-1056365.  
 PR 30-MAR-2001; 2001DE-1016220.  
 XX  
 PA (IPFP-) IPF PHARM GMBH.  
 XX  
 PI Forssmann W, Kluever E, Conejo-Garcia J, Adermann K, Bals R;  
 PI Maegert H;  
 XX  
 DR WPI; 2002-435959/46.  
 XX  
 PT New human beta-defensin 3, useful for treating or preventing microbial  
 PT infection and tumors, also related nucleic acid -  
 XX  
 PS Examples; Page 15; 36pp; German.  
 XX  
 CC The present invention relates to human beta-defensin-3 (hBD-3) and its  
 CC derivatives. The peptide, its coding sequence and vectors containing the  
 CC coding sequence are useful in (gene) therapy and diagnosis, especially  
 CC for preventing or treating a wide range of microbial infections  
 CC (particularly Burkholderia cepacia and Pseudomonas aeruginosa in the  
 CC respiratory tract, especially in cases of cystic fibrosis, and  
 CC Helicobacter pylori, also inflammatory diseases of the gastrointestinal  
 CC and urogenital tracts, sepsis and yeast infections), and for inducing  
 CC apoptosis for treating malignant melanoma and tumours. The present  
 CC sequence is a fragment of human BD-3.  
 XX  
 SQ Sequence 17 AA;  
  
 Query Match 62.5%; Score 5; DB 23; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
 |||||  
 Db 9 RGGRC 13

RESULT 14  
 AAW36214  
 ID AAW36214 standard; peptide; 18 AA.  
 XX  
 AC AAW36214;  
 XX  
 DT 12-FEB-1998 (first entry)  
 XX  
 DE Antimicrobial protegrin peptide (14).  
 XX  
 KW Antimicrobial protegrin; broad spectrum; gram-positive; bacterium;  
 KW gram-negative; bacteria; yeast; fungus; fungi; protozoa; virus;

KW retrovirus; HIV; human immunodeficiency virus; preservation;  
 KW disinfection; prophylaxis; treatment; infection; disease;  
 KW conjunctivitis; keratitis; corneal ulcer; stomach ulcer; STD;  
 KW Helicobacter pylori; sexually transmitted disease; oral mucositis;  
 KW gram-negative sepsis; endocarditis; pneumonia; biocidal; biostatic;  
 KW respiratory infection; urinary tract infection; MRSA; protozoan;  
 KW vancomycin resistant Enterococcus; pathogen; multi-drug resistance;  
 KW penicillin resistant Streptococcus pneumoniae; pig; porcine;  
 KW methicillin resistant Staphylococcus aureus; systemic candidiasis.

XX

OS Synthetic.

OS Sus scrofa.

XX

FH	Key	Location/Qualifiers
FT	Misc-difference	1
FT		/note= "D-form residue"
FT	Misc-difference	2
FT		/note= "D-form residue"
FT	Misc-difference	3
FT		/note= "D-form residue"
FT	Misc-difference	4
FT		/note= "D-form residue"
FT	Misc-difference	5
FT		/note= "D-form residue"
FT	Misc-difference	6
FT		/note= "D-form residue"
FT	Misc-difference	7
FT		/note= "D-form residue"
FT	Misc-difference	8
FT		/note= "D-form residue"
FT	Misc-difference	9
FT		/note= "D-form residue"
FT	Misc-difference	10
FT		/note= "D-form residue"
FT	Misc-difference	11
FT		/note= "D-form residue"
FT	Misc-difference	12
FT		/note= "D-form residue"
FT	Misc-difference	13
FT		/note= "D-form residue"
FT	Misc-difference	14
FT		/note= "D-form residue"
FT	Misc-difference	15
FT		/note= "D-form residue"
FT	Misc-difference	16
FT		/note= "D-form residue"
FT	Misc-difference	17
FT		/note= "D-form residue"
FT	Misc-difference	18
FT		/note= "D-form residue"

XX

PN W09718826-A1.

XX

PD 29-MAY-1997.

XX

PF 22-NOV-1996; 96WO-US18544.

XX



PR 21-NOV-1996; 96US-0752852.  
PR 22-NOV-1995; 95US-0562346.  
PR 17-MAY-1996; 96US-0649811.  
PR 01-AUG-1996; 96US-0690921.

XX

PA (INTR-) INTRABIOTICS PHARM INC.

PA (REGC ) UNIV CALIFORNIA.

XX

PI Chang CC, Chen J, Gu CL, Lehrer RI, Steinberg DA;

XX

DR WPI; 1997-297871/27.

XX

PT New antimicrobial protegrin peptide(s) - having activity against  
PT bacteria, yeast, fungi, protozoa and certain strains of viruses  
PT (e.g. HIV)

XX

PS Claim 23; Page 104; 130pp; English.

XX

CC The present sequence is an antimicrobial protegrin peptide, which  
CC has a broad spectrum of activity against microbial targets,  
CC including gram-positive and gram-negative bacteria, yeast, fungi,  
CC protozoa and certain strains of viruses and retroviruses, e.g. HIV.  
CC It can be used to preserve or disinfect a variety of materials,  
CC including medical equipment, foodstuffs, cosmetics, contact lens  
CC solutions, medicaments or other nutrient containing materials. It  
CC can also be used for the prophylaxis or treatment of microbial  
CC infections or diseases in plants and animals, e.g. conjunctivitis,  
CC keratitis, corneal ulcers, stomach ulcers associated with  
CC Helicobacter pylori, sexually transmitted diseases, gram-negative  
CC sepsis, endocarditis, pneumonia and other respiratory infections,  
CC urinary tract infections, systemic candidiasis and oral mucositis.  
CC It is biostatic or biocidal against clinically relevant pathogens  
CC exhibiting multi-drug resistance, e.g. vancomycin resistant  
CC Enterococcus faecium or faecalis, penicillin resistant  
CC Streptococcus pneumoniae and methicillin resistant Staphylococcus  
CC aureus (MRSA). It is given at a dosage of 0.1 to 5, preferably 0.5  
CC to 1 mg/kg/day, by injection.

XX

SQ Sequence 18 AA;

Query Match 62.5%; Score 5; DB 18; Length 18;

Best Local Similarity 100.0%; Pred. No. 96;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LRGGR 7

||||

Db 14 LRGGR 18

RESULT 15

AAW23726

ID AAW23726 standard; peptide; 18 AA.

XX

AC AAW23726;

XX

DT 16-OCT-1997 (first entry)

XX

DE Tachytegrin and/or parevin compound peptide.  
XX  
KW Protegrin; tachyplesin; antimicrobial; antiviral; oral mucositis;  
KW sexually transmitted disease; endotoxin; HIV; Staphylococcus aureus;  
KW Pseudomonas; H. pylori; Chlamydia trachomatis; Treponema pallidum;  
KW Neisseria gonorrhoeae; Trichomonas vaginalis; Herpes simplex type 2;  
KW animal; plant.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 5..8  
FT /label= Optional  
FT Disulfide-bond 5..16  
FT /label= Optional  
FT Disulfide-bond 8..13  
FT /label= Optional  
FT Disulfide-bond 13..16  
FT /label= Optional  
FT Modified-site 18  
FT /note= "Optionally amidated"  
XX  
PN WO9702287-A1.  
XX  
PD 23-JAN-1997.  
XX  
PF 03-JUL-1996; 96WO-US11323.  
XX  
PR 06-JUL-1995; 95US-0000898.  
XX  
PA (INTR-) INTRABIOTICS PHARM INC.  
XX  
PI Chang CC, Gu CL, Harwig SSL, Lehrer RI;  
XX  
DR WPI; 1997-108914/10.  
XX  
PT Cpd. based on protegrin(s) and tachyplesin(s) - designated  
PT tachytegrin(s) and parevin(s), useful as antimicrobial and antiviral  
PT agents, e.g. for treatment of sexually transmitted diseases  
XX  
PS Claim 21; Page 42; 59pp; English.  
XX  
CC The present sequence represents a peptide sequence found in compounds  
CC based on protegrins and tachyplesins, designated tachytegrins and  
CC parevins. The present sequence is from a parevin as the peptide has Cys  
CC residues at positions 5 and 16, where tachytegrins have Cys residues at  
CC positions 4 and 17. Both compounds can be used as preservatives (e.g.  
CC in foodstuffs, cosmetics or medicaments) or in treatment of viral or  
CC microbial infection in animals or plants. They are useful, e.g. in  
CC treatment of oral mucositis, infection by Staphylococcus aureus,  
CC Pseudomonas or H. pylori, or sexually transmitted diseases, including  
CC those caused by Chlamydia trachomatis, Treponema pallidum, Neisseria  
CC gonorrhoeae, Trichomonas vaginalis, Herpes simplex type 2 and HIV. DNA  
CC encoding the compounds can also be expressed in situ, in animals or  
CC plants, to combat infections. The compounds can also be used as  
CC standards in antimicrobial assays and in binding endotoxins.  
XX

SQ Sequence 18 AA;

Query Match 62.5%; Score 5; DB 18; Length 18;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

Search completed: November 13, 2003, 10:32:53  
Job time : 26.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:38:27 ; Search time 15.6667 Seconds  
(without alignments)  
93.222 Million cell updates/sec

Title: US-09-228-866-4  
Perfect score: 8  
Sequence: 1 CVLRGGRC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 124183

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	62.5	9	9	US-09-834-765-241	Sequence 241, Appl
2	5	62.5	9	9	US-09-834-765-354	Sequence 354, Appl
3	5	62.5	10	9	US-09-834-765-298	Sequence 298, Appl
4	5	62.5	10	9	US-09-834-765-399	Sequence 399, Appl
5	5	62.5	16	10	US-09-865-943-61	Sequence 61, Appl
6	5	62.5	16	10	US-09-865-943-63	Sequence 63, Appl
7	5	62.5	16	10	US-09-865-943-186	Sequence 186, Appl
8	5	62.5	16	10	US-09-865-943-188	Sequence 188, Appl
9	5	62.5	17	10	US-09-865-943-13	Sequence 13, Appl
10	5	62.5	17	10	US-09-865-943-21	Sequence 21, Appl
11	5	62.5	17	10	US-09-865-943-47	Sequence 47, Appl
12	5	62.5	17	10	US-09-865-943-93	Sequence 93, Appl
13	5	62.5	17	10	US-09-865-943-95	Sequence 95, Appl
14	5	62.5	17	10	US-09-865-943-105	Sequence 105, Appl
15	5	62.5	17	10	US-09-865-943-109	Sequence 109, Appl
16	5	62.5	17	10	US-09-865-943-122	Sequence 122, Appl
17	5	62.5	17	10	US-09-865-943-126	Sequence 126, Appl
18	5	62.5	17	10	US-09-865-943-130	Sequence 130, Appl
19	5	62.5	17	10	US-09-865-943-143	Sequence 143, Appl
20	5	62.5	18	10	US-09-865-943-1	Sequence 1, Appli
21	5	62.5	18	10	US-09-865-943-2	Sequence 2, Appli
22	5	62.5	18	10	US-09-865-943-4	Sequence 4, Appli
23	5	62.5	18	10	US-09-865-943-5	Sequence 5, Appli
24	5	62.5	18	10	US-09-865-943-11	Sequence 11, Appl
25	5	62.5	18	10	US-09-865-943-17	Sequence 17, Appl
26	5	62.5	18	10	US-09-865-943-19	Sequence 19, Appl
27	5	62.5	18	10	US-09-865-943-29	Sequence 29, Appl
28	5	62.5	18	10	US-09-865-943-31	Sequence 31, Appl
29	5	62.5	18	10	US-09-865-943-33	Sequence 33, Appl
30	5	62.5	18	10	US-09-865-943-45	Sequence 45, Appl
31	5	62.5	18	10	US-09-865-943-51	Sequence 51, Appl
32	5	62.5	18	10	US-09-865-943-60	Sequence 60, Appl
33	5	62.5	18	10	US-09-865-943-68	Sequence 68, Appl
34	5	62.5	18	10	US-09-865-943-69	Sequence 69, Appl
35	5	62.5	18	10	US-09-865-943-70	Sequence 70, Appl
36	5	62.5	18	10	US-09-865-943-76	Sequence 76, Appl
37	5	62.5	18	10	US-09-865-943-86	Sequence 86, Appl
38	5	62.5	18	10	US-09-865-943-88	Sequence 88, Appl
39	5	62.5	18	10	US-09-865-943-104	Sequence 104, App
40	5	62.5	18	10	US-09-865-943-107	Sequence 107, App
41	5	62.5	18	10	US-09-865-943-108	Sequence 108, App
42	5	62.5	18	10	US-09-865-943-113	Sequence 113, App
43	5	62.5	18	10	US-09-865-943-114	Sequence 114, App

44	5	62.5	18	10	US-09-865-943-115	Sequence 115, App
45	5	62.5	18	10	US-09-865-943-121	Sequence 121, App

# ALIGNMENTS

## RESULT 1

```

US-09-834-765-241
; Sequence 241, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 241
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-241

```

```

Query Match          62.5%; Score 5; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 LRGGR 7
        |||||
Db      5 LRGGR 9

```

## RESULT 2

```

US-09-834-765-354
; Sequence 354, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765

```

; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 354  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-354

Query Match 62.5%; Score 5; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGGR 7  
|||  
Db 5 LRGGR 9

RESULT 3

US-09-834-765-298  
; Sequence 298, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 298  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-298

Query Match 62.5%; Score 5; DB 9; Length 10;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGGR 7  
|||  
Db 5 LRGGR 9

RESULT 4

US-09-834-765-399

```

; Sequence 399, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 399
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-399

```

```

Query Match          62.5%; Score 5; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 60;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 LRGGR 7
        |||||
Db      6 LRGGR 10

```

```

RESULT 5
US-09-865-943-61
; Sequence 61, Application US/09865943
; Patent No. US20020147301A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/865,943
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/128,344
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 16
; TYPE: PRT

```

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-865-943-61

Query Match 62.5%; Score 5; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 6

US-09-865-943-63  
; Sequence 63, Application US/09865943  
; Patent No. US20020147301A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/865,943  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 09/128,344  
; PRIOR FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-865-943-63

Query Match 62.5%; Score 5; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||||  
Db 1 RGGRC 5

RESULT 7

US-09-865-943-186  
; Sequence 186, Application US/09865943  
; Patent No. US20020147301A1  
; GENERAL INFORMATION:



```

; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/865,943
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/128,344
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: DISULFID
; LOCATION: (5)...(16)
US-09-865-943-186

```

```

Query Match          62.5%;  Score 5;  DB 10;  Length 16;
Best Local Similarity 100.0%;  Pred. No. 86;
Matches      5;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      4 RGGRC 8
        |||||
Db      1 RGGRC 5

```

```

RESULT 8
US-09-865-943-188
; Sequence 188, Application US/09865943
; Patent No. US20020147301A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/865,943
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/128,344
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 188

```

; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
; NAME/KEY: DISULFID  
; LOCATION: (5)...(16)  
US-09-865-943-188

Query Match 62.5%; Score 5; DB 10; Length 16;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||  
Db 1 RGGRC 5

RESULT 9

US-09-865-943-13  
; Sequence 13, Application US/09865943  
; Patent No. US20020147301A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/865,943  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 09/128,344  
; PRIOR FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-865-943-13

Query Match 62.5%; Score 5; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||  
Db 1 RGGRC 5

RESULT 10

US-09-865-943-21  
 ; Sequence 21, Application US/09865943  
 ; Patent No. US20020147301A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lehrer, Robert I.  
 ; APPLICANT: Harwig, Sylvia S. L.  
 ; APPLICANT: Chang, Conway C.  
 ; APPLICANT: Gu, Chee L.  
 ; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
 ; FILE REFERENCE: 8067-0053-999  
 ; CURRENT APPLICATION NUMBER: US/09/865,943  
 ; CURRENT FILING DATE: 2001-05-24  
 ; PRIOR APPLICATION NUMBER: US 09/128,344  
 ; PRIOR FILING DATE: 1998-08-03  
 ; PRIOR APPLICATION NUMBER: US 08/647,622  
 ; PRIOR FILING DATE: 1996-07-03  
 ; PRIOR APPLICATION NUMBER: US 60/000,898  
 ; PRIOR FILING DATE: 1995-07-06  
 ; NUMBER OF SEQ ID NOS: 201  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 21  
 ; LENGTH: 17  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic polypeptide  
 US-09-865-943-21

Query Match 62.5%; Score 5; DB 10; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
 |||||  
 Db 1 RGGRC 5

RESULT 11  
 US-09-865-943-47  
 ; Sequence 47, Application US/09865943  
 ; Patent No. US20020147301A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lehrer, Robert I.  
 ; APPLICANT: Harwig, Sylvia S. L.  
 ; APPLICANT: Chang, Conway C.  
 ; APPLICANT: Gu, Chee L.  
 ; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
 ; FILE REFERENCE: 8067-0053-999  
 ; CURRENT APPLICATION NUMBER: US/09/865,943  
 ; CURRENT FILING DATE: 2001-05-24  
 ; PRIOR APPLICATION NUMBER: US 09/128,344  
 ; PRIOR FILING DATE: 1998-08-03  
 ; PRIOR APPLICATION NUMBER: US 08/647,622  
 ; PRIOR FILING DATE: 1996-07-03  
 ; PRIOR APPLICATION NUMBER: US 60/000,898  
 ; PRIOR FILING DATE: 1995-07-06  
 ; NUMBER OF SEQ ID NOS: 201

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: MOD_RES
; LOCATION: (1)...(17)
; OTHER INFORMATION: All genetically encoded amino acids are in the
; OTHER INFORMATION: D-configuration
US-09-865-943-47
```

```
Query Match          62.5%; Score 5; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 90;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      4 RGGRC 8
        |||||
Db      1 RGGRC 5
```

# RESULT 12

```
US-09-865-943-93
; Sequence 93, Application US/09865943
; Patent No. US20020147301A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/865,943
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/128,344
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-865-943-93
```

```
Query Match          62.5%; Score 5; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 90;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      4 RGGRC 8
        |||||
```

Db 1 RGGRC 5

RESULT 13

US-09-865-943-95  
; Sequence 95, Application US/09865943  
; Patent No. US20020147301A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/865,943  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 09/128,344  
; PRIOR FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622  
; PRIOR FILING DATE: 1996-07-03  
; PRIOR APPLICATION NUMBER: US 60/000,898  
; PRIOR FILING DATE: 1995-07-06  
; NUMBER OF SEQ ID NOS: 201  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 95  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic polypeptide  
US-09-865-943-95

Query Match 62.5%; Score 5; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
|||  
Db 1 RGGRC 5

RESULT 14

US-09-865-943-105  
; Sequence 105, Application US/09865943  
; Patent No. US20020147301A1  
; GENERAL INFORMATION:  
; APPLICANT: Lehrer, Robert I.  
; APPLICANT: Harwig, Sylvia S. L.  
; APPLICANT: Chang, Conway C.  
; APPLICANT: Gu, Chee L.  
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS  
; FILE REFERENCE: 8067-0053-999  
; CURRENT APPLICATION NUMBER: US/09/865,943  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 09/128,344  
; PRIOR FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 08/647,622

```

; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: DISULFID
; LOCATION: (5)...(16)
; NAME/KEY: DISULFID
; LOCATION: (8)...(13)
US-09-865-943-105

```

```

Query Match          62.5%; Score 5; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 90;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 RGGRC 8
        |||||
Db      1 RGGRC 5

```

# RESULT 15

```

US-09-865-943-109
; Sequence 109, Application US/09865943
; Patent No. US20020147301A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia S. L.
; APPLICANT: Chang, Conway C.
; APPLICANT: Gu, Chee L.
; TITLE OF INVENTION: PAREVINS AND TACHYTEGRINS
; FILE REFERENCE: 8067-0053-999
; CURRENT APPLICATION NUMBER: US/09/865,943
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/128,344
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: US 08/647,622
; PRIOR FILING DATE: 1996-07-03
; PRIOR APPLICATION NUMBER: US 60/000,898
; PRIOR FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; NAME/KEY: DISULFID
; LOCATION: (5)...(16)
; NAME/KEY: DISULFID
; LOCATION: (8)...(13)

```

US-09-865-943-109

Query Match 62.5%; Score 5; DB 10; Length 17;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGRC 8  
    |||||  
Db 1 RGGRC 5

Search completed: November 13, 2003, 11:12:32  
Job time : 15.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:26:01 ; Search time 7.83333 Seconds  
(without alignments)  
98.215 Million cell updates/sec

Title: US-09-228-866-4  
Perfect score: 8  
Sequence: 1 CVLRGGRC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3752

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pirl:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
-----						

1	4	50.0	12	2	S43170	kinesin light chain
2	4	50.0	16	2	E49655	T-cell-receptor beta
3	4	50.0	17	2	A60570	Ig mu heavy chain
4	4	50.0	18	2	A43334	orf1 5' of aadR -
5	4	50.0	19	2	S43641	carboxylesterase (
6	3	37.5	7	2	PT0667	T-cell receptor beta
7	3	37.5	7	4	S15597	orf 4 rara 5'-regi
8	3	37.5	8	2	PT0279	Ig heavy chain CRD
9	3	37.5	9	2	F28854	fibrinopeptide B -
10	3	37.5	9	2	D24180	fibrinogen beta chain
11	3	37.5	9	2	B45020	probable minipolypeptide
12	3	37.5	9	2	PT0670	T-cell receptor beta
13	3	37.5	9	2	PT0634	T-cell receptor beta
14	3	37.5	10	2	A61131	hydrin 2 - bullfro
15	3	37.5	10	2	S06964	hypothetical protein
16	3	37.5	10	2	PT0245	Ig heavy chain CRD
17	3	37.5	11	2	PH0947	T-cell receptor beta
18	3	37.5	12	2	A33900	hydrin 1 - African
19	3	37.5	12	2	PH1606	Ig H chain V-D-J region
20	3	37.5	12	2	PH0930	T-cell receptor beta
21	3	37.5	13	2	PH0805	T-cell receptor alpha
22	3	37.5	14	1	NTKN1M	alpha-conotoxin MI
23	3	37.5	14	2	A58963	alpha-conotoxin Cn
24	3	37.5	14	2	JH0328	probursin tetradec
25	3	37.5	14	2	S58426	spermadhesin AWN homolog
26	3	37.5	14	2	PH0804	T-cell receptor alpha
27	3	37.5	15	2	S24159	leukocyte elastase
28	3	37.5	15	2	I49420	placental lactogen
29	3	37.5	15	2	PH0136	T-cell receptor beta
30	3	37.5	15	2	PH0764	T-cell receptor beta
31	3	37.5	15	2	S71396	dihydropyrimidine
32	3	37.5	16	2	A60551	leukocyte elastase
33	3	37.5	16	2	A49226	major outer membrane
34	3	37.5	16	2	PT0296	Ig heavy chain CDR
35	3	37.5	16	2	PH1346	Ig heavy chain DJ
36	3	37.5	16	2	S57517	T cell receptor beta
37	3	37.5	16	2	PH1453	T-cell receptor alpha
38	3	37.5	17	2	PN0587	tyrosine 3-monooxygenase
39	3	37.5	17	2	B20242	pyruvate kinase (E)
40	3	37.5	17	2	PT0234	Ig heavy chain CRD
41	3	37.5	17	2	S51736	T-cell receptor beta
42	3	37.5	18	2	S28424	34K protein - rape
43	3	37.5	19	2	S32548	glutathione transferase
44	3	37.5	19	2	A48354	nonstructural protein
45	3	37.5	19	2	I49037	TcR delta chain V-

#### ALIGNMENTS

##### RESULT 1

S43170

kinesin light chain - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C;Accession: S43170

R;Chernajovsky, Y.; Brown, A.; Jones, T.A.; Sheer, D.



submitted to the EMBL Data Library, December 1992  
A;Description: Promoter first exon/intron characterization and chromosomal location of the human light chain (beta) kinesin gene.  
A;Reference number: S43170  
A;Accession: S43170  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-12 <CHE>  
A;Cross-references: EMBL:X69658; NID:g468786; PIDN:CAA49349.1; PID:g468787

Query Match 50.0%; Score 4; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGG 6  
    ||||  
Db 2 LRGG 5

#### RESULT 2

E49655  
T-cell-receptor beta chain variable region, TCR V beta (clone 3) - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 30-May-1997  
C;Accession: E49655  
R;Grom, A.A.; Thompson, S.D.; Luyrink, L.; Passo, M.; Choi, E.; Glass, D.N. Proc. Natl. Acad. Sci. U.S.A. 90, 11104-11108, 1993  
A;Title: Dominant T-cell-receptor beta chain variable region V beta 14+ clones in juvenile rheumatoid arthritis.  
A;Reference number: A49655; MUID:94068553; PMID:8248215  
A;Accession: E49655  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: nucleic acid  
A;Residues: 1-16 <GRO>  
A;Experimental source: peripheral blood lymphocytes  
A;Note: sequence extracted from NCBI backbone (NCBIP:140450)  
C;Keywords: T-cell receptor

Query Match 50.0%; Score 4; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGRC 8  
    ||||  
Db 10 GGRC 13

#### RESULT 3

A60570  
Ig mu heavy chain disease extra sequence - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 16-Aug-1996  
C;Accession: A60570  
R;Mihaesco, C.; Ferrara, P.; Guillemot, J.C.; Congy, N.; Gendron, M.C.; Roy, J.P.; Sizaret, P.Y.; Mihaesco, E.  
Mol. Immunol. 27, 771-776, 1990

A;Title: A new extra sequence at the amino terminal of a mu heavy chain disease protein (DAG).

A;Reference number: A60570; MUID:90384486; PMID:2119480

A;Accession: A60570

A;Molecule type: protein

A;Residues: 1-17 <MIH>

C;Comment: This sequence is derived from the amino terminus of an abnormal immunoglobulin found in lymphoproliferative malignancies.

C;Keywords: immunoglobulin

Query Match 50.0%; Score 4; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLRG 5  
|||  
Db 7 VLRG 10

#### RESULT 4

A43334

orf1 5' of aadR - Rhodopseudomonas palustris (fragment)

C;Species: Rhodopseudomonas palustris

C;Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Oct-1999

C;Accession: A43334

R;Dispensa, M.; Thomas, C.T.; Kim, M.K.; Perrotta, J.A.; Gibson, J.; Harwood, C.S.

J. Bacteriol. 174, 5803-5813, 1992

A;Title: Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators.

A;Reference number: A43334; MUID:92394882; PMID:1522059

A;Accession: A43334

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-18 <DIS>

A;Cross-references: GB:M92426; NID:g151870; PIDN:AAA26089.1; PID:g151871

A;Note: sequence extracted from NCBI backbone (NCBIN:112964, NCBIP:112965)

Query Match 50.0%; Score 4; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RGGR 7  
|||  
Db 2 RGGR 5

#### RESULT 5

S43641

carboxylesterase (EC 3.1.1.1), thermostable - Alicyclobacillus acidocaldarius (fragment)

C;Species: Alicyclobacillus acidocaldarius

C;Date: 19-Mar-1997 #sequence\_revision 01-Aug-1997 #text\_change 17-Mar-1999

C;Accession: S43641

R;Manco, G.; di Gennaro, S.; de Rosa, M.; Rossi, M.

Eur. J. Biochem. 221, 965-972, 1994

A;Title: Purification and characterization of a thermostable carboxylesterase from the thermoacidophilic eubacterium *Bacillus acidocaldarius*.  
A;Reference number: S43641; MUID:94237161; PMID:8181479  
A;Accession: S43641  
A;Molecule type: protein  
A;Residues: 1-19 <MAN>  
A;Experimental source: strain MT3  
A;Note: the source is designated as *Bacillus acidocaldarius*  
C;Keywords: carboxylic ester hydrolase; monomer

Query Match 50.0%; Score 4; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLRG 5  
|||  
Db 13 VLRG 16

#### RESULT 6

PT0667

T-cell receptor beta chain V-D-J region (121-21) - mouse (fragment)

C;Species: *Mus musculus* (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0667

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0667

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-7 <FEE>

A;Experimental source: day 4 postnatal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 37.5%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
|||  
Db 5 GGR 7

#### RESULT 7

S15597

orf 4 rara 5'-region - human

C;Species: *Homo sapiens* (man)

C;Date: 04-Jun-1999 #sequence\_revision 04-Jun-1999 #text\_change 28-Jun-1999

C;Accession: S15597

R;Brand, N.J.; Petkovich, M.; Chambon, P.

Nucleic Acids Res. 18, 6799-6806, 1990

A;Title: Characterization of a functional promoter for the human retinoic acid receptor-alpha (hRAR-alpha).

A;Reference number: S15594; MUID:91088249; PMID:2175878

A;Accession: S15597  
A;Molecule type: DNA  
A;Residues: 1-7 <BRA>  
A;Cross-references: EMBL:X56058; NID:g35876  
A;Note: this ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0  
C;Comment: This sequence is not thought to be translated.  
C;Genetics:  
A;Gene: GDB:RARA  
A;Cross-references: GDB:120337; OMIM:180240  
A;Map position: 17q12-17q12

Query Match 37.5%; Score 3; DB 4; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGG 6  
|||  
Db 3 RGG 5

#### RESULT 8

PT0279

Ig heavy chain CRD3 region (clone 4-91A) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C;Accession: PT0279

R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G. J. Exp. Med. 173, 395-407, 1991

A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes.

A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0279

A;Molecule type: DNA

A;Residues: 1-8 <YAM>

A;Experimental source: B lymphocyte

C;Keywords: heterotetramer; immunoglobulin

Query Match 37.5%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGG 6  
|||  
Db 6 RGG 8

#### RESULT 9

F28854

fibrinopeptide B - gelada baboon

C;Species: Theropithecus gelada (gelada baboon)

C;Date: 19-May-1989 #sequence\_revision 19-May-1989 #text\_change 18-Aug-2000

C;Accession: F28854

R;Nakamura, S.; Takenaka, O.; Takahashi, K. J. Biochem. 94, 1973-1978, 1983

A;Title: Fibrinopeptides A and B of baboons (*Papio anubis*, *Papio hamadryas*, and *Theropithecus gelada*): their amino acid sequences and evolutionary rates and a molecular phylogeny for the baboons.

A;Reference number: A91973; MUID:84161822; PMID:6423621

A;Accession: F28854

A;Molecule type: protein

A;Residues: 1-9 <NAK>

C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide ring homology

Query Match 37.5%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
|||  
Db 7 GGR 9

#### RESULT 10

D24180

fibrinogen beta chain - red guenon (fragment)

N;Contains: fibrinopeptide B

C;Species: *Erythrocebus patas* (red guenon, hussar)

C;Date: 05-Jun-1988 #sequence\_revision 10-Mar-1994 #text\_change 26-Jan-1996

C;Accession: D24180

R;Nakamura, S.; Takenaka, O.; Takahashi, K.

J. Biochem. 97, 1487-1492, 1985

A;Title: Fibrinopeptides A and B of Japanese monkey (*Macaca fuscata*) and patas monkey (*Erythrocebus patas*): their amino acid sequences, restricted mutations, and a molecular phylogeny for macaques, guenons, and baboons.

A;Reference number: A91990; MUID:85289140; PMID:3928610

A;Accession: D24180

A;Molecule type: protein

A;Residues: 1-9 <NAK>

C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide ring homology

Query Match 37.5%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
|||  
Db 7 GGR 9

#### RESULT 11

B45020

probable minipolypeptide - mouse

C;Species: *Mus musculus* (house mouse)

C;Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Aug-2000

C;Accession: B45020

R;Nielsen, A.L.; Pallisgaard, N.; Pedersen, F.S.; Jorgensen, P.

Mol. Cell. Biol. 12, 3449-3459, 1992

A;Title: Murine helix-loop-helix transcriptional activator proteins binding to the E-box motif of the Akv murine leukemia virus enhancer identified by cDNA cloning.

A;Reference number: A45020; MUID:92334344; PMID:1321336

A;Accession: B45020

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-9 <NIE>

A;Experimental source: NIH 3T3 fibroblasts

A;Note: sequence extracted from NCBI backbone (NCBIN:108710, NCBIIP:108711)

C;Superfamily: unassigned animal peptides

Query Match 37.5%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7

|||

Db 5 GGR 7

#### RESULT 12

PT0670

T-cell receptor beta chain V-D-J region (121-1BN) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0670

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0670

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-9 <FEE>

A;Experimental source: day 4 postnatal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 37.5%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7

|||

Db 6 GGR 8

#### RESULT 13

PT0634

T-cell receptor beta chain V-D-J region (121-2CL) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0634

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0634

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-9 <FEE>

A;Experimental source: newborn thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 37.5%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
|||  
Db 7 GGR 9

#### RESULT 14

A61131

hydrin 2 - bullfrog

N;Alternate names: AVT-related peptide

N;Contains: Arg-vasotocin

C;Species: Rana catesbeiana (bullfrog)

C;Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 11-Jul-1997

C;Accession: A61131

R;Iwamuro, S.; Hayashi, H.; Yamashita, M.; Kikuyama, S.

Gen. Comp. Endocrinol. 84, 412-418, 1991

A;Title: Arginine vasotocin (AVT) and AVT-related peptide are major aldosterone-releasing factors in the bullfrog neurointermediate lobe.

A;Reference number: A61131; MUID:92225310; PMID:1808022

A;Accession: A61131

A;Molecule type: protein

A;Residues: 1-10 <IWA>

C;Comment: Both the decapeptide (hydrin 2) and the amidated nonapeptide (Arg-vasotocin) stimulate the release of aldosterone and thus have an antidiuretic effect.

C;Superfamily: oxytocin-neurophysin

C;Keywords: amidated carboxyl end; neuropeptide; pituitary

F;1-9/Product: Arg-vasotocin #status experimental <RVAS>

F;9/Modified site: amidated carboxyl end (Gly) (amide in mature form from following glycine) #status experimental

Query Match 37.5%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGG 6  
|||  
Db 8 RGG 10

#### RESULT 15

S06964

hypothetical protein (nifA 5' region) - Rhizobium leguminosarum (fragment)

C;Species: Rhizobium leguminosarum

C;Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 08-Oct-1999  
C;Accession: S06964  
R;Roelvink, P.W.; Hontelez, J.G.J.; van Kammen, A.; van den Bos, R.C.  
Mol. Microbiol. 3, 1441-1447, 1989  
A;Title: Nucleotide sequence of the regulatory nifA gene of Rhizobium  
leguminosarum PRE: transcriptional control sites and expression in Escherichia  
coli.  
A;Reference number: S06964; MUID:90136072; PMID:2693897  
A;Accession: S06964  
A;Molecule type: DNA  
A;Residues: 1-10 <ROE>  
A;Cross-references: EMBL:X17073; NID:g46208; PIDN:CAA34923.1; PID:g809748

Query Match 37.5%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGG 6  
|||  
Db 1 RGG 3

Search completed: November 13, 2003, 10:39:54  
Job time : 8.83333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:55:06 ; Search time 4.33333 Seconds  
(without alignments)  
86.819 Million cell updates/sec

Title: US-09-228-866-4  
Perfect score: 8  
Sequence: 1 CVLRGGRC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1246

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,



and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	4	50.0	18	1	YAA5_RHOPA	Q02005	rhodopseudo
2	3	37.5	8	1	COW2_CONPU	P58785	conus purpu
3	3	37.5	9	1	FAR6_MACRS	P83279	macrobrachi
4	3	37.5	9	1	FIBB_ERYPA	P19346	erythrocebu
5	3	37.5	9	1	FIBB_THEGE	P19342	theropithec
6	3	37.5	12	1	UKA2_HUMAN	P31144	homo sapien
7	3	37.5	12	1	V14K_WSSV	P82006	white spot
8	3	37.5	14	1	CXA1_CONCN	P56973	conus conso
9	3	37.5	14	1	CXA1_CONMA	P01521	conus magus
10	3	37.5	14	1	FIBB_MANLE	P14474	mandrillus
11	3	37.5	14	1	SAP2_ARBPU	P11760	arbacia pun
12	3	37.5	16	1	BRB_BASAL	P83187	basella alb
13	3	37.5	20	1	CISY_STRHY	P20903	streptomyce
14	3	37.5	20	1	OMPW_VIBAL	P83151	vibrio algi
15	3	37.5	20	1	PORC_METTM	P80902	methanobact
16	3	37.5	20	1	PUFK_RHOSH	Q53121	rhodobacter
17	2	25.0	7	1	CARP_MYTED	P10420	mytilus edu
18	2	25.0	7	1	FAR1_HELTI	P41871	helisoma tr
19	2	25.0	7	1	FAR1_MACRS	P83274	macrobrachi
20	2	25.0	7	1	FAR1_PROCL	P38499	procambarus
21	2	25.0	7	1	FAR2_ASCSU	P31890	ascaris suu
22	2	25.0	7	1	FAR2_PROCL	P38498	procambarus
23	2	25.0	7	1	FAR5_HIRME	P42564	hirudo medi
24	2	25.0	7	1	UF04_MOUSE	P38642	mus musculu
25	2	25.0	8	1	AL16_CARMA	P81819	carcinus ma
26	2	25.0	8	1	ALL5_CYDPO	P82156	cydia pomon
27	2	25.0	8	1	ALL9_CARMA	P81812	carcinus ma
28	2	25.0	8	1	CAD1_ENTFA	P13268	enterococcu
29	2	25.0	8	1	FAR1_PANRE	P41872	panagrellus
30	2	25.0	8	1	FAR1_PENMO	P83316	penaeus mon
31	2	25.0	8	1	FAR2_MACRS	P83275	macrobrachi
32	2	25.0	8	1	FAR3_HOMAM	P41486	homarus ame
33	2	25.0	8	1	FAR4_HOMAM	P41487	homarus ame
34	2	25.0	8	1	FAR4_MACRS	P83277	macrobrachi
35	2	25.0	8	1	UF06_MOUSE	P38644	mus musculu
36	2	25.0	8	1	WP1_PERAT	P83195	perkinsus a
37	2	25.0	9	1	BS43_SERPL	P83375	serratia pl
38	2	25.0	9	1	CONO_CONST	P05487	conus stria
39	2	25.0	9	1	D1_NEPNO	P24816	nephrops no
40	2	25.0	9	1	DNF1_LOCFI	P16339	locusta mig
41	2	25.0	9	1	DSIP_RABIT	P01158	oryctolagus
42	2	25.0	9	1	FAR2_PANRE	P41873	panagrellus
43	2	25.0	9	1	FAR3_MACRS	P83276	macrobrachi
44	2	25.0	9	1	FAR3_PENMO	P83318	penaeus mon
45	2	25.0	9	1	FAR4_PENMO	P83319	penaeus mon

# ALIGNMENTS

# RESULT 1

## YAA5\_RHOPA

ID YAA5\_RHOPA STANDARD; PRT; 18 AA.  
AC Q02005;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein in aadR 5'region (Fragment).  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CGA009;  
RX MEDLINE=92394882; PubMed=1522059;  
RA Dispensa M., Thomas C.T., Kim M.K., Perrotta J.A., Gibson J.,  
RA Harwood C.S.;  
RT "Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate  
RT is dependent on AadR, a member of the cyclic AMP receptor protein  
RT family of transcriptional regulators.";  
RL J. Bacteriol. 174:5803-5813(1992).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M92426; AAA26089.1; -.  
DR PIR; A43334; A43334.  
KW Hypothetical protein.  
FT NON TER 1 1  
SQ SEQUENCE 18 AA; 2173 MW; 61DABDFE4E317B0A CRC64;

Query Match 50.0%; Score 4; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGR 7  
| | | |  
Db 2 RGGR 5

# RESULT 2

## COW2\_CONPU

ID COW2\_CONPU STANDARD; PRT; 8 AA.  
AC P58785;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leu-contryphan-P.  
OS Conus purpurascens (Purple cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;

OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=41690;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.  
 RC STRAIN=Clipperton Island; TISSUE=Venom;  
 RX MEDLINE=99388839; PubMed=10461743;  
 RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,  
 RA Olivera B.M.;  
 RT "A novel D-leucine-containing Conus peptide: diverse conformational  
 RT dynamics in the contryphan family."  
 RL J. Pept. Res. 54:93-99(1999).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -!- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.  
 CC -!- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.  
 KW Toxin; Hydroxylation; D-amino acid.  
 FT DISULFID 2 8  
 FT MOD\_RES 4 4 D-LEUCINE.  
 SQ SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CVL 3  
 |||  
 Db 2 CVL 4

# RESULT 3

## FAR6\_MACRS

ID FAR6\_MACRS STANDARD; PRT; 9 AA.  
 AC P83279;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE FMRFamide-like neuropeptide FLP6 (DGGRNFLRF-amide).  
 OS Macrobrachium rosenbergii (Giant fresh water prawn).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
 OC Palaemonoidea; Palaemonidae; Macrobrachium.  
 OX NCBI\_TaxID=79674;  
 RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RC TISSUE=Eyestalk;  
 RX MEDLINE=21107394; PubMed=11179812;  
 RA Sithigorngul P., Saraithongkum W., Longyant S., Panchan N.,  
 RA Sithigorngul W., Petsom A.;  
 RT "Three more novel FMRFamide-like neuropeptide sequences from the  
 RT eyestalk of the giant freshwater prawn Macrobrachium rosenbergii."  
 RL Peptides 22:191-197(2001).  
 CC -!- MASS SPECTROMETRY: MW=1080.7; METHOD=MALDI.  
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 DR GO; GO:0007218; P:neuropeptide signaling pathway; IDA.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 9 9 AMIDATION.

SQ SEQUENCE 9 AA; 1081 MW; 26800729C4540878 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
|||  
Db 2 GGR 4

RESULT 4

FIBB\_ERYPA

ID FIBB\_ERYPA STANDARD; PRT; 9 AA.  
AC P19346;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Erythrocebus patas (Red guenon) (Hussar).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Erythrocebus.  
OX NCBI\_TaxID=9538;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=85289140; PubMed=3928610;  
RA Nakamura S., Takenaka O., Takahashi K.;  
RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and  
RT patas monkey (Erythrocebus patas): their amino acid sequences,  
RT restricted mutations, and a molecular phylogeny for macaques,  
RT guenons, and baboons.";  
RL J. Biochem. 97:1487-1492(1985).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR PIR; D24180; D24180.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma.  
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 1020 MW; 69FE7879C732CB1B CRC64;

Query Match 37.5%; Score 3; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
|||

Db

7 GGR 9

RESULT 5

FIBB\_THEGE

ID FIBB\_THEGE STANDARD; PRT; 9 AA.  
AC P19342;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Theropithecus gelada (Gelada baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Theropithecus.  
OX NCBI\_TaxID=9565;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=84161822; PubMed=6423621;  
RA Nakamura S., Takenaka O., Takahashi K.;  
RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,  
RT and Theropithecus gelada): their amino acid sequences and  
RT evolutionary rates and a molecular phylogeny for the baboons.";  
RL J. Biochem. 94:1973-1978(1983).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR PIR; F28854; F28854.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma.  
FT PEPTIDE 1 9 FIBRINOPEPTIDE B.  
FT NON\_TER 9 9  
SQ SEQUENCE 9 AA; 977 MW; DDFE7879C7287B06 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GGR 7

|||

Db 7 GGR 9

RESULT 6

UKA2\_HUMAN

ID UKA2\_HUMAN STANDARD; PRT; 12 AA.  
AC P31144;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of epidermal keratinocytes (Spot 1617)  
 DE (Fragments).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Keratinocytes;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
 RA Vandekerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RT protein database of normal human epidermal keratinocytes.";  
 RL Electrophoresis 13:960-969(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 6.93, ITS MW IS: 81.6 kDa.  
 DR Aarhus/Ghent-2DPAGE; 1617; IEF.  
 FT NON\_TER 1 1  
 FT NON\_CONS 7 8  
 FT UNSURE 8 8  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1351 MW; D6CD4A5E75F2C1F6 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GGR 7  
 |||  
 Db 10 GGR 12

#### RESULT 7

V14K\_WSSV  
 ID V14K\_WSSV STANDARD; PRT; 12 AA.  
 AC P82006;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 14.5 kDa structural polyprotein (Fragment).  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.  
 OX NCBI\_TaxID=92652;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=South Carolina;  
 RX MEDLINE=20214217; PubMed=10752552;  
 RA Wang Q., Poulos B.T., Lightner D.V.;  
 RT "Protein analysis of geographic isolates of shrimp white spot syndrome  
 RT virus.";  
 RL Arch. Virol. 145:263-274(2000).  
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF THE VIRION.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1242 MW; 24B8DE4FFD21A338 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 7.7e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGG 6  
|||  
Db 3 RGG 5

RESULT 8

CXA1\_CONCN

ID CXA1\_CONCN STANDARD; PRT; 14 AA.  
AC P56973;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Alpha-conotoxin CnIA [Contains: Alpha-conotoxin CnIB].  
OS Conus consors (Singed cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=101297;  
RN [1]  
RP SEQUENCE, SYNTHESIS, CHARACTERIZATION, AND STRUCTURE BY NMR.  
RC TISSUE=Venom;  
RX MEDLINE=99255390; PubMed=10320362;  
RA Favreau P., Krimm I., le Gall F., Bobenrieth M.J., Lamthanh H.,  
RA Bouet F., Servent D., Molgo J., Menez A., Letourneux Y.,  
RA Lancelin J.-M.;  
RT "Biochemical characterization and nuclear magnetic resonance  
RT structure of novel alpha-conotoxins isolated from the venom of Conus  
RT consors.";  
RL Biochemistry 38:6317-6326(1999).  
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
CC INHIBIT THEM. THIS PEPTIDE SEEMS TO BE A POTENT AND SELECTIVE  
CC BLOCKER OF MUSCULAR SUBTYPE OF NACHR.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE  
CC FAMILY.  
DR PIR; A58963; A58963.  
DR PDB; 1B45; 09-JUL-99.  
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;  
KW Acetylcholine receptor inhibitor; Amidation; 3D-structure.  
FT PEPTIDE 1 14 ALPHA-CONOTOXIN CNIA.  
FT PEPTIDE 3 14 ALPHA-CONOTOXIN CNIB.  
FT DISULFID 3 8  
FT DISULFID 4 14  
FT MOD\_RES 14 14 AMIDATION.  
FT HELIX 6 8  
FT TURN 9 10  
SQ SEQUENCE 14 AA; 1548 MW; DEEE91969BF5E5BD CRC64;

Query Match 37.5%; Score 3; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GRC 8  
 |||  
Db 1 GRC 3

RESULT 9

CXA1\_CONMA

ID CXA1\_CONMA STANDARD; PRT; 14 AA.  
AC P01521;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Alpha-conotoxin MI (M1).  
OS Conus magus (Magus cone).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxID=6492;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=83073458; PubMed=7149738;  
RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;  
RT "Isolation and structure of a peptide toxin from the marine snail  
RT Conus magus.";  
RL Arch. Biochem. Biophys. 218:329-334(1982).  
RN [2]  
RP DISULFIDE BONDS.  
RX MEDLINE=84032400; PubMed=6630187;  
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;  
RT "Conotoxin MI. Disulfide bonding and conformational states.";  
RL J. Biol. Chem. 258:12247-12251(1983).  
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY  
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS  
CC INHIBIT THEM.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE  
CC FAMILY.  
DR PIR; A01784; NTKN1M.  
DR HSSP; P56973; 1B45.  
KW Postsynaptic neurotoxin; Neurotoxin; Toxin;  
KW Acetylcholine receptor inhibitor; Amidation.  
FT DISULFID 3 8  
FT DISULFID 4 14  
FT MOD\_RES 14 14 AMIDATION.  
SQ SEQUENCE 14 AA; 1499 MW; DEEE91898BF5E5BD CRC64;

Query Match 37.5%; Score 3; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GRC 8  
 |||  
Db 1 GRC 3



# RESULT 10

## FIBB\_MANLE

ID FIBB\_MANLE STANDARD; PRT; 14 AA.  
AC P14474;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
GN FGB.  
OS Mandrillus leucophaeus (Drill) (Papio leucophaeus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
OC Cercopithecinae; Mandrillus.  
OX NCBI\_TaxID=9568;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=69115139; PubMed=4974768;  
RA Doolittle R.F., Glasgow C., Mross G.A.;  
RT "Characterization of fibrinopeptides A and B from a drill (Mandrillus  
RT leucophaeus).";  
RL Biochim. Biophys. Acta 175:217-219(1969).  
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
CC AGGREGATION.  
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
DR InterPro; IPR002181; Fibrinogen\_C.  
DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
KW Blood coagulation; Plasma; Pyrrolidone carboxylic acid.  
FT PEPTIDE 1 14 FIBRINOPEPTIDE B.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT NON\_TER 14 14  
SQ SEQUENCE 14 AA; 1434 MW; 6695B0F11EF72E1B CRC64;

Query Match 37.5%; Score 3; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
| | |  
Db 12 GGR 14

# RESULT 11

## SAP2\_ARBPU

ID SAP2\_ARBPU STANDARD; PRT; 14 AA.  
AC P11760;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 01-MAR-1992 (Rel. 21, Last annotation update)  
DE Resact (Sperm-activating peptide) (SAP-IIA).  
OS Arbacia punctulata (Punctuate sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinoidea; Euechinoidea; Echinacea; Arbacoida; Arbaciidae; Arbacia.  
 OX NCBI\_TaxID=7641;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Egg;  
 RX MEDLINE=85054981; PubMed=6150045;  
 RA Suzuki N., Shimomura H., Radany E.W., Ramarao C.S., Ward G.E.,  
 RA Bentley J.K., Garbers D.L.;  
 RT "A peptide associated with eggs causes a mobility shift in a major  
 RT plasma membrane protein of spermatozoa."  
 RL J. Biol. Chem. 259:14874-14879(1984).  
 RN [2]  
 RP DISULFIDE BOND.  
 RX MEDLINE=92097763; PubMed=1756858;  
 RA Yoshimo K.-I., Takao T., Shimonishi Y., Suzuki N.;  
 RT "Determination of the amino acid sequence of an intramolecular  
 RT disulfide linkage-containing sperm-activating peptide by tandem mass  
 RT spectrometry."  
 RL FEBS Lett. 294:179-182(1991).  
 CC -!- FUNCTION: CAUSE STIMULATION OF SPERM RESPIRATION AND MOTILITY  
 CC THROUGH INTRACELLULAR ALKALINIZATION, TRANSIENT ELEVATIONS OF  
 CC CAMP, CGMP AND CLACIUM LEVELS IN SPERM CELLS, AND TRANSIENT  
 CC ACTIVATION AND SUBSEQUENT INACTIVATION OF THE MEMBRANE FORM OF  
 CC GUANYLATE CYCLASE.  
 CC -!- SIMILARITY: SMALL TO S.PURPURATUS SPERACT.  
 KW Amidation.  
 FT DISULFID 1 8  
 FT MOD\_RES 14 14 AMIDATION.  
 SQ SEQUENCE 14 AA; 1246 MW; 39745AA33EBE41B8 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GGR 7  
 |||  
 Db 11 GGR 13

# RESULT 12

## BRB\_BASAL

ID BRB\_BASAL STANDARD; PRT; 16 AA.  
 AC P83187;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Beta-basrubin (Fragment).  
 OS Basella alba (Malabar spinach) (Ceylon spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllidae; Caryophyllales; Basellaceae; Basella.  
 OX NCBI\_TaxID=3589;  
 RN [1]  
 RP SEQUENCE, AND FUNCTION.  
 RC TISSUE=Seed;  
 RX MEDLINE=21547763; PubMed=11688973;  
 RA Wang H., Ng T.B.;

RT "Novel antifungal peptides from ceylon spinach seeds.";  
 RL Biochem. Biophys. Res. Commun. 288:765-770(2001).  
 CC -!- FUNCTION: Possesses antifungal activity against B.cinerea,  
 CC M.arachidicola and F.oxysporum but not C.comatus and R.solani.  
 CC Inhibits HIV-1 reverse transcriptase and cell-free translation.  
 DR GO; GO:0003799; F:antifungal peptide activity; IDA.  
 DR GO; GO:0017148; P:negative regulation of protein biosynthesis; IDA.  
 KW Fungicide.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1952 MW; 28F9FE4FC181682C CRC64;

Query Match 37.5%; Score 3; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 9.8e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRG 5  
 |||  
 Db 13 LRG 15

# RESULT 13

## CISY\_STRHY

ID CISY\_STRHY STANDARD; PRT; 20 AA.  
 AC P20903;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Citrate synthase (EC 2.3.3.1) (Fragment).  
 GN GLTA.  
 OS Streptomyces hygroscopicus.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1912;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=SF1293;  
 RX MEDLINE=90334852; PubMed=1368511;  
 RA Shimotohno K.W., Imai S., Murakami T., Seto H.;  
 RT "Purification and characterization of citrate synthase from  
 RT Streptomyces hygroscopicus SF-1293 and comparison of its properties  
 RT with those of 2-phosphinomethylmalic acid synthase.";  
 RL Agric. Biol. Chem. 54:463-470(1990).  
 CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + oxaloacetate = citrate +  
 CC CoA.  
 CC -!- ENZYME REGULATION: ALLOSTERICALLY INHIBITED BY NADH.  
 CC -!- PATHWAY: Tricarboxylic acid cycle.  
 CC -!- SUBUNIT: Homohexamer.  
 CC -!- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
 CC CAPABLE OF OXIDATIVE METABOLISM.  
 CC -!- SIMILARITY: Belongs to the citrate synthase family.  
 DR PIR; PQ0046; PQ0046.  
 DR InterPro; IPR002020; Citrate\_synt.  
 DR PROSITE; PS00480; CITRATE\_SYNTHASE; PARTIAL.  
 KW Transferase; Tricarboxylic acid cycle; Allosteric enzyme.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2234 MW; C527EC7A87119597 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLR 4  
|||  
Db 6 VLR 8

RESULT 14

OMPW\_VIBAL

ID OMPW\_VIBAL STANDARD; PRT; 20 AA.  
AC P83151;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Outer membrane protein W (Outer membrane protein 25Va) (Omp25Va)  
DE (Fragment).  
GN OMPW.  
OS *Vibrio alginolyticus*.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=663;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=NCIMB 1903T;  
RA Onji M., Hirabayashi J., Suzuki S.;  
RT "Characterization of major outer membrane proteins of *Vibrio*  
RT *alginolyticus* and the stability against proteases.";  
RL Microbes Environ. 0:0-0(2002).  
CC -!- SUBCELLULAR LOCATION: Outer membrane.  
CC -!- SIMILARITY: BELONGS TO THE OMPW/ALKL FAMILY.  
KW Outer membrane.  
FT NON\_TER 20 20  
SQ SEQUENCE 20 AA; 2096 MW; D29EE7FCA16C0D37 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VLR 4  
|||  
Db 7 VLR 9

RESULT 15

PORC\_METTM

ID PORC\_METTM STANDARD; PRT; 20 AA.  
AC P80902;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Pyruvate synthase subunit PORC (EC 1.2.7.1) (Pyruvate oxidoreductase  
DE gamma chain) (POR) (Pyruvic-ferredoxin oxidoreductase gamma subunit)  
DE (Fragment).  
GN PORC.  
OS *Methanobacterium thermoautotrophicum* (strain Marburg / DSM 2133).

OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=79929;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=97261844; PubMed=9108258;  
 RA Tersteegen A., Linder D., Thauer R.K., Hedderich R.;  
 RT "Structures and functions of four anabolic 2-oxoacid oxidoreductases  
 RT in Methanobacterium thermoautotrophicum.";  
 RL Eur. J. Biochem. 244:862-868(1997).  
 CC -!- CATALYTIC ACTIVITY: Pyruvate + CoA + oxidized ferredoxin = acetyl-  
 CC CoA + CO(2) + reduced ferredoxin.  
 CC -!- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (BY SIMILARITY).  
 CC -!- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE  
 CC GAMMA CHAIN.  
 CC -!- MISCELLANEOUS: As a pH optimum of 10.0 and an optimal temperature  
 CC of 80 degrees Celsius.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.  
 KW Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S.  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2167 MW; 481532134D42F353 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGG 6  
 |||  
 Db 9 RGG 11

Search completed: November 13, 2003, 10:33:58  
 Job time : 5.33333 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:58:36 ; Search time 19.6667 Seconds  
 (without alignments)  
 104.971 Million cell updates/sec

Title: US-09-228-866-4  
 Perfect score: 8  
 Sequence: 1 CVLRGGRC 8

Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4	50.0	12	2	Q46712	Q46712 escherichia
2	4	50.0	12	4	Q13865	Q13865 homo sapien
3	4	50.0	12	10	Q93X21	Q93x21 zea mays (m
4	4	50.0	14	15	Q8UT17	Q8ut17 human immun
5	4	50.0	18	11	Q9QVJ3	Q9qvj3 mus sp. car
6	4	50.0	19	3	Q9UR83	Q9ur83 saccharomyc
7	4	50.0	19	4	Q8IU87	Q8iu87 homo sapien
8	4	50.0	21	4	Q8NET0	Q8net0 homo sapien
9	4	50.0	21	5	Q9U5M8	Q9u5m8 megaselia s
10	3	37.5	7	11	Q63480	Q63480 rattus norv
11	3	37.5	8	12	O89965	O89965 polyomaviru
12	3	37.5	9	11	Q8R514	Q8r514 rattus norv
13	3	37.5	9	11	Q99JF4	Q99jf4 mus musculu
14	3	37.5	9	12	Q9IBM8	Q9ibm8 simian viru
15	3	37.5	9	12	Q9PYK1	Q9pyk1 simian viru
16	3	37.5	10	2	Q52837	Q52837 rhizobium l
17	3	37.5	10	4	Q96QA7	Q96qa7 homo sapien
18	3	37.5	10	4	Q9H1I5	Q9h1i5 homo sapien
19	3	37.5	10	4	Q8NEY9	Q8ney9 homo sapien
20	3	37.5	10	4	Q9UCU6	Q9ucu6 homo sapien
21	3	37.5	10	5	P82383	P82383 drosophila
22	3	37.5	10	12	O90349	O90349 hepatitis g
23	3	37.5	10	12	O90346	O90346 hepatitis g
24	3	37.5	11	3	Q9HFN8	Q9hfn8 candida rug

25	3	37.5	11	4	Q9UNL8	Q9unl8 homo sapien
26	3	37.5	11	13	Q9PS22	Q9ps22 xenopus lae
27	3	37.5	12	4	Q9H1Z6	Q9hlz6 homo sapien
28	3	37.5	12	6	O46664	O46664 macropus ro
29	3	37.5	12	6	Q9XT42	Q9xt42 canis famil
30	3	37.5	13	2	Q9REI2	Q9rei2 acidiphiliu
31	3	37.5	13	7	Q9TNQ8	Q9tnq8 homo sapien
32	3	37.5	13	8	Q8WGC2	Q8wgc2 isocheles p
33	3	37.5	14	2	Q9R2Z4	Q9r2z4 campylobact
34	3	37.5	14	2	Q9R2Z5	Q9r2z5 campylobact
35	3	37.5	14	2	Q9X715	Q9x715 campylobact
36	3	37.5	14	2	Q9WW79	Q9ww79 campylobact
37	3	37.5	14	10	Q9XGM4	Q9xgm4 arabidopsis
38	3	37.5	14	12	Q86578	Q86578 sigma virus
39	3	37.5	14	15	Q8JDM3	Q8jdm3 human immun
40	3	37.5	14	15	Q8JDM7	Q8jdm7 human immun
41	3	37.5	14	15	Q8JDM0	Q8jdm0 human immun
42	3	37.5	15	1	Q9UWG1	Q9uwg1 pyrococcus,
43	3	37.5	15	2	Q9R598	Q9r598 micrococcus
44	3	37.5	15	4	Q9Y4Z9	Q9y4z9 homo sapien
45	3	37.5	15	5	Q9TWF3	Q9twf3 dictyosteli

#### ALIGNMENTS

##### RESULT 1

Q46712

ID Q46712 PRELIMINARY; PRT; 12 AA.  
 AC Q46712;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
 DE Hypothetical 1.3 kDa protein.  
 OS Escherichia coli.  
 OG Plasmid R100.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86250614; PubMed=3522549;  
 RA Fee B.E., Dempsey W.B.;  
 RT "Cloning, mapping, and sequencing of plasmid R100 traM and finP  
 RT genes.";  
 RL J. Bacteriol. 167:336-345(1986).  
 DR EMBL; M13054; AAA88503.1; -.  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 12 AA; 1295 MW; 640674B567A775A4 CRC64;

Query Match 50.0%; Score 4; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 RGGR 7  
 ||||  
 Db 2 RGGR 5

# RESULT 2

Q13865

ID Q13865 PRELIMINARY; PRT; 12 AA.  
AC Q13865;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Kinesin light chain.  
GN BETA-KINESIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=97101110; PubMed=8945637;  
RA Chernajovsky Y., Brown A., Jones T.A.;  
RT "Promoter first exon/intron characterization and chromosomal location  
RT of the human light chain (beta) kinesin gene."  
RL DNA Cell Biol. 15:965-974(1996).  
DR EMBL; X69658; CAA49349.1; -.  
SQ SEQUENCE 12 AA; 1274 MW; 3FB8D34EE165A5B8 CRC64;

Query Match 50.0%; Score 4; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LRGG 6  
| | | |  
Db 2 LRGG 5

# RESULT 3

Q93X21

ID Q93X21 PRELIMINARY; PRT; 12 AA.  
AC Q93X21;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Polyubiquitin homolog (Fragment).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. INRA 258; TISSUE=Leaf;  
RX MEDLINE=96236829; PubMed=8680303;  
RA Didierjean L., Frendo P., Nasser W., Genot G., Marivet J., Burkard G.;  
RT "Heavy-metal-responsive genes in maize: identification and comparison  
RT of their expression upon various forms of abiotic stress."  
RL Planta 199:1-8(1996).  
DR EMBL; S82313; AAB47175.1; -.